

特許協力条約

PCT

特許性に関する国際予備報告（特許協力条約第二章）

（法第 12 条、法施行規則第 56 条）
〔PCT36 条及び PCT 規則 70〕

REC'D 22 DEC 2005

WIPO

PCT

出願人又は代理人 の書類記号 CO-F03092-00	今後の手続きについては、様式 PCT/IPEA/416 を参照すること。	
国際出願番号 PCT/JP2004/015263	国際出願日 (日.月.年) 15.10.2004	優先日 (日.月.年) 15.12.2003
国際特許分類 (IPC) Int.Cl. G06F17/28		
出願人（氏名又は名称） 有限会社言語技術研究所		

1. この報告書は、PCT35 条に基づきこの国際予備審査機関で作成された国際予備審査報告である。
法施行規則第 57 条（PCT36 条）の規定に従い送付する。

2. この国際予備審査報告は、この表紙を含めて全部で 4 ページからなる。

3. この報告には次の附属物件も添付されている。

a. 附属書類は全部で 53 ページである。

補正されて、この報告の基礎とされた及び／又はこの国際予備審査機関が認めた訂正を含む明細書、請求の範囲及び／又は図面の用紙（PCT 規則 70.16 及び実施細則第 607 号参照）

第 I 欄 4. 及び補充欄に示したように、出願時における国際出願の開示の範囲を超えた補正を含むものとこの国際予備審査機関が認定した差替え用紙

b. 電子媒体は全部で _____ (電子媒体の種類、数を示す)。
配列表に関する補充欄に示すように、電子形式による配列表又は配列表に関連するテーブルを含む。
(実施細則第 802 号参照)

4. この国際予備審査報告は、次の内容を含む。

第 I 欄 国際予備審査報告の基礎
 第 II 欄 優先権
 第 III 欄 新規性、進歩性又は産業上の利用可能性についての国際予備審査報告の不作成
 第 IV 欄 発明の單一性の欠如
 第 V 欄 PCT35 条(2)に規定する新規性、進歩性又は産業上の利用可能性についての見解、それを裏付けるための文献及び説明
 第 VI 欄 ある種の引用文献
 第 VII 欄 国際出願の不備
 第 VIII 欄 国際出願に対する意見

国際予備審査の請求書を受理した日 14.10.2005	国際予備審査報告を作成した日 28.11.2005
名称及びあて先 日本国特許庁 (IPEA/JP) 郵便番号 100-8915 東京都千代田区霞が関三丁目 4 番 3 号	特許庁審査官（権限のある職員） 和田 財太 電話番号 03-3581-1101 内線 3599 5M 9459

第I欄 報告の基礎

1. 言語に関し、この予備審査報告は以下のものを基礎とした。

出願時の言語による国際出願

出願時の言語から次の目的のための言語である _____ 語に翻訳された、この国際出願の翻訳文

國際調査 (PCT規則12.3(a)及び23.1(b))

國際公開 (PCT規則12.4(a))

國際予備審査 (PCT規則55.2(a)又は55.3(a))

2. この報告は下記の出願書類を基礎とした。(法第6条 (PCT14条) の規定に基づく命令に応答するために提出された差替え用紙は、この報告において「出願時」とし、この報告に添付していない。)

出願時の国際出願書類

明細書

第 1-2, 16, 21, 22, 26, 29, 30, 34, 36, 38- ページ、出願時に提出されたもの
43, 45-49

第 3, 3/1, 4-14, 14/1-14/8, 15, 15/1, 17 ページ*, 14. 10. 2005

付けで国際予備審査機関が受理したもの

-20, 23-25, 27, 27/1, 28, 28/1, 31, 32

, 33, 33/1, 35, 37, 44, 44/1

第 _____ ページ*, _____

付けで国際予備審査機関が受理したもの

請求の範囲

第 _____ 項、出願時に提出されたもの

第 _____ 項*, PCT 19条の規定に基づき補正されたもの

第 1-19

項*, 14. 10. 2005

付けで国際予備審査機関が受理したもの

第 _____ 項*

項*

付けで国際予備審査機関が受理したもの

図面

第 1-5 ページ/図、出願時に提出されたもの

第 _____ ページ/図*, _____

付けで国際予備審査機関が受理したもの

第 _____ ページ/図*

付けで国際予備審査機関が受理したもの

配列表又は関連するテーブル

配列表に関する補充欄を参照すること。

3. 補正により、下記の書類が削除された。

明細書 第 _____ ページ

請求の範囲 第 _____ 項

図面 第 _____ ページ/図

配列表 (具体的に記載すること)

配列表に関するテーブル (具体的に記載すること) _____

4. この報告は、補充欄に示したように、この報告に添付されかつ以下に示した補正が出願時における開示の範囲を超えてされたものと認められるので、その補正がされなかったものとして作成した。(PCT規則 70.2(c))

明細書 第 _____ ページ

請求の範囲 第 _____ 項

図面 第 _____ ページ/図

配列表 (具体的に記載すること)

配列表に関するテーブル (具体的に記載すること) _____

* 4. に該当する場合、その用紙に "superseded" と記入されることがある。

第III欄 新規性、進歩性又は産業上の利用可能性についての見解の不作成

次に関して、当該請求の範囲に記載されている発明の新規性、進歩性又は産業上の利用可能性につき、次の理由により審査しない。

国際出願全体

請求の範囲 5-9, 12-14, 17-19

理由：

この国際出願又は請求の範囲 _____ は、国際予備審査をすることを要しない次の事項を内容としている（具体的に記載すること）。

明細書、請求の範囲若しくは図面（次に示す部分）又は請求の範囲 _____ の記載が、不明確であるため、見解を示すことができない（具体的に記載すること）。

全部の請求の範囲又は請求の範囲 _____ が、明細書による十分な裏付けを欠くため、見解を示すことができない（具体的に記載すること）。

請求の範囲 5-9, 12-14, 17-19 について、国際調査報告が作成されていない。

入手可能な配列表が存在せず、有意義な見解を示すことができなかった。

出願人は所定の期間内に、

実施細則の附属書Cに定める基準を満たす紙形式の配列表を提出しなかったため、国際予備審査機関は、認められた形式及び方法で配列表を入手することができなかった。

実施細則の附属書Cに定める基準を満たす電子形式の配列表を提出しなかったため、国際予備審査機関は、認められた形式及び方法で配列表を入手することができなかった。

PCT規則13の3.1(a)又は(b)及び13の3.2に基づく命令に応じた、要求された配列表の遅延提出手数料を支払わなかった。

入手可能な配列表に関連するテーブルが存在しないため、有意義な見解を示すことができなかった。すなわち、出願人が、所定の期間内に、実施細則の附属書Cの2に定める技術的な要件を満たす電子形式のテーブルを提出しなかったため、国際予備審査機関は、認められた形式及び方法でテーブルを入手することができなかった。

ヌクレオチド又はアミノ酸の配列表に関連するテーブルが電子形式のみで提出された場合において、当該テーブルが、実施細則の附属書Cの2に定める技術的な要件を満たしていない。

詳細については補充欄を参照すること。

第V欄 新規性、進歩性又は産業上の利用可能性についての法第12条(PCT35条(2))に定める見解、それを裏付ける文献及び説明

1. 見解

新規性 (N)	請求の範囲 1-4, 10, 11, 15-16	有
	請求の範囲	無
進歩性 (I S)	請求の範囲 1-4, 10, 11, 15-16	有
	請求の範囲	無
産業上の利用可能性 (I A)	請求の範囲 1-4, 10, 11, 15-16	有
	請求の範囲	無

2. 文献及び説明 (PCT規則70.7)

文献1 : JP 2003-263434 A (株式会社国際電気通信基礎技術研究所) 2003.09.19,
【0021】-【0022】(ファミリーなし)

文献2 : JP 04-160473 A (株式会社日立製作所) 1992.06.03,
第4頁左下欄第4行-第4頁右下欄第6行 (ファミリーなし)

文献3 : JP 2000-020524 A (富士ゼロックス株式会社) 2000.01.21,
【0008】(ファミリーなし)

請求項1-4, 10, 11, 15-16に係る発明は、国際調査報告に引用されたいずれの文献にも記載されておらず、当業者にとって、自明なものでもない。

語句自体に付随する背景知識、人が母語で話したり書いたりするための文書を作る時に思い浮かべる単位（すなわちチャンク）を踏まえた上で、原文の対訳を求めるようすれば、目的言語の文として自然な翻訳文が得られることになるが、文脈の判断や語句自体に付随する背景知識の判断は容易でない上に、チャンクもその区切りが曖昧であり、機械翻訳において、チャンク単位で原文の対訳を求めて翻訳文を得ることは実現できていないのが実情であった。

[0009] 本発明は上記事実を考慮して成されたもので、原言語の原文から目的言語の文として自然な対訳文が得られる可能性の高い、原文中の翻訳対象語句の適切な対訳を得ることができる対訳判断装置、対訳判断方法及びプログラムを得ることが目的である。

課題を解決するための手段

[0010] 上記目的を達成するために請求項1記載の発明に係る対訳判断装置は、複数の単語から成る原言語の自然文を目的言語の対訳文と対応付けて複数記憶する記憶手段と、前記記憶手段に記憶されている原言語の複数の自然文の中から、原言語の原文中の翻訳対象語句が含まれている自然文を検索する検索手段と、前記検索手段による検索によって抽出された自然文に基づいて、前記翻訳対象語句と原言語の同一の文中に出現する頻度の高い頻出語句を認識する認識手段と、前記検索手段による検索によって抽出された自然文のうち、前記認識手段によって認識されかつ前記原文中に存在する特定の頻出語句と、前記翻訳対象語句が各々含まれている自然文の対訳文を参照することで、前記翻訳対象語句及び前記特定の頻出語句が各々含まれる自然文の対訳文における翻訳対象語句についての高頻度の対訳を認識し、認識した高頻度の対訳を、前記原文中の前記翻訳対象語句の対訳と判断する対訳判断手段と、を含んで構成されている。

[0011] 請求項1記載の発明では、複数の単語から成る原言語の自然文を目的言語の対訳文と対応付けて記憶手段に複数記憶している。なお、本発明に係る自然文は、従来の機械翻訳における辞書のように、単語単位での分割や多義語の抽出等の編集・加工を行っていない文、文節、連語、定型表現、コロケーションの少なくとも何れかであり、この自然文を目的言語の対訳文と対応付けて記憶することで、例えば単語単位で辞書に登録する場合のように、多義語に多数の対訳を対応付ける必要が無く

なる（多義語に多数の対訳を対応付けた場合、対訳を選択する必要が生じ、選択誤りが生ずる可能性がある）と共に、自然文に対応する対訳文を目的言語の文として自然な文にすることができる。なお、本発明において、自然文以外に単語とその対訳も

記憶手段に記憶させても構わない。

[0012] また、請求項1記載の発明では、記憶手段に記憶されている原言語の複数の自然文の中から、原言語の原文（文、文節、連語、定型表現、コロケーションの何れでもよい）中の翻訳対象語句が含まれている自然文が検索手段によって検索される。翻訳対象語句は原文のうち特に翻訳すべき語句であり、単語であってもよいし、複数の単語で構成されていてもよい。例えば本発明に係る対訳判断装置を人間（翻訳者）が翻訳する際の電子辞書として用いる場合、翻訳対象語句は翻訳者によって指定される。また、例えば本発明に係る対訳判断装置を機械翻訳装置又は自動通訳装置の一部として用いる場合、翻訳対象語句は、本発明に係る対訳判断装置によって判断された対訳を利用して機械翻訳を行う機械翻訳装置又は自動通訳装置によって指定される。また、翻訳対象語句を含む原文についても、利用者（翻訳者等）に指定させるようにしてもよいし、自動的に判断する（例えば翻訳対象語句を含む文又は文節を自動的に原文と判断する等）ようにしてもよい。検索手段は、翻訳対象語句が含まれている自然文を検索するので、この検索により、対応する対訳文の中に翻訳対象語句の対訳が含まれている自然文が抽出されることになる。

[0013] ここで、請求項1記載の発明では、検索手段による検索によって抽出された自然文に基づいて、翻訳対象語句と原言語の同一の文中に出現する頻度の高い頻出語句が認識手段によって認識され、対訳判断手段は、検索手段による検索によって抽出された自然文のうち、認識手段によって認識されかつ原文中に存在する特定の頻出語句と、翻訳対象語句が各々含まれている自然文の対訳文を参照することで、翻訳対象語句及び特定の頻出語句が各々含まれる自然文の対訳文における翻訳対象語句についての高頻度の対訳を認識し、認識した高頻度の対訳を、原文中の翻訳対象語句の対訳と判断している。

[0014] 翻訳対象語句と原言語の同一の文中に出現する頻度（コリレーションという）の高い特定の頻出語句が原文中に存在している場合、翻訳対象語句の対訳は、翻訳対象語句及び特定の頻出語句が各々含まれている自然文の対訳文における翻訳対象語句の対訳に一致している可能性が高い。但し、翻訳対象語句及び特定の頻出語句が各々含まれている自然文の対訳文における翻訳対象語句の適切な対訳は同じである可能性が高いものの、記憶手段に記憶されている上記の自然文の中に、翻訳対象語

句の適切な対訳が相違している自然文が混在している可能性もある。上記に基づき請求項1記載の発明では、翻訳対象語句とのコリレーションの高い頻出語句を認識し、認識した頻出語句のうち原文中に存在する特定の頻出語句と、翻訳対象語句が各々含まれている自然文の対訳文を参照することで、翻訳対象語句及び特定の頻出語句が各々含まれる自然文の対訳文における翻訳対象語句についての高頻度の対訳を認識し、認識した高頻度の対訳を、原文中の翻訳対象語句の対訳と判断するので、頻出語句に基づいて、原言語の原文から目的言語の文として自然な対訳文が得られる可能性の高い、原文中の翻訳対象語句の適切な対訳（原文から目的言語の文として自然な対訳文が得られる可能性の高い対訳）を得ることができる。

[0015] 訳文選択の際の前後の文脈の判断は、状況に応じて無数に存在し、事前に特定することは困難であるが、請求項1記載の発明では、同じ文章の中で同時に出現する確率の高い語句に着目し、その語句と翻訳対象語句が同時に出現する自然文の対訳文を参照することで、翻訳対象語句についての対訳を判断するので、原文の前後の文脈を認識することなく、結果として多義語についても、おおよそ前後の文脈を踏まえた上で適切な対訳を得ることができる。なお、上記の頻出語句の認識は、例えば原言語の同一の文中に出現する頻度の高い語句同士をテーブルに登録しておき、該テーブルを参照することで行うことも可能であるが、請求項1記載の発明では、検索手段による検索によって抽出された自然文に基づいて頻出語句を認識しているので、上記のテーブル作成の手間が省けると共に、テーブルを記憶するために必要となる記憶容量を節減できる、という効果も得られる。

[0016] 請求項2記載の発明に係る対訳判断装置は、複数の単語から成る原言語の自然文を目的言語の対訳文と対応付けて複数記憶する記憶手段と、前記記憶手段に記憶されている原言語の複数の自然文の中から、原言語の原文中の翻訳対象語句が含まれている自然文を検索する検索手段と、前記原文中に存在しつつ前記検索手段による検索によって抽出された自然文中に含まれていない着目語句と代替可能な代替語句を判断する判断手段と、前記検索手段による検索によって抽出された自然文のうち、前記判断手段によって判断された代替語句及び前記翻訳対象語句が各々含まれている自然文の対訳文における少なくとも前記翻訳対象語句の対訳を、前記原文中の少なくとも前記翻訳対象語句の対訳と判断する対訳判断手段と、を含んで構成され

ている。

[0017] 請求項2記載の発明では、請求項1記載の発明と同様に、複数の単語から成る原言語の自然文が目的言語の対訳文と対応付けて記憶手段に複数記憶されており、検索手段により、記憶手段に記憶されている原言語の複数の自然文の中から、原言語の原文中の翻訳対象語句が含まれている自然文が検索される。ここで、検索手段による検索により、原文中の特定の語句が別の語句に入れ替わった自然文が抽出されることがあるが、このとき原文の意味と自然文の意味が類似であれば、特定の語句と別の語句は代替可能な関係にあると考えられる。そして、検索手段による検索によって複数の自然文が複数抽出された場合にも、その中に原文に対して特定の語句が代替可能な関係にある別の語句（代替語句）に入れ替わっている自然文が存在している場合には、この自然文は原文と意味が類似である可能性が高いので、この自然文を選択することが望ましい。

[0018] 上記に基づき請求項2記載の発明では、原文中に存在しかつ検索手段による検索によって抽出された自然文中に含まれていない着目語句と代替可能な代替語句が判断手段によって判断され、対訳判断手段は、検索手段による検索によって抽出された自然文のうち、判断手段によって判断された代替語句及び翻訳対象語句が各々含まれている自然文の対訳文における少なくとも翻訳対象語句の対訳を、原文中の少なくとも翻訳対象語句の対訳と判断している。これにより、代替語句に基づいて、原言語の原文から目的言語の文として自然な対訳文が得られる可能性の高い、原文中の翻訳対象語句の適切な対訳（原文から目的言語の文として自然な対訳文が得られる可能性の高い対訳）を得ることができる。

[0019] なお、請求項2記載の発明に係る判断手段による代替語句の判断は、例えば代替可能な関係にある語句同士をテーブルに登録しておき、このテーブルを参照することで行うようにしてもよいが、例えば請求項3に記載したように、着目語句を含む自然文を記憶手段に記憶されている複数の自然文の中から検索し、該検索によって抽出された自然文と同一の構文の自然文を記憶手段に記憶されている自然文の中から検索し、該検索によって抽出された自然文中で着目語句と置き換わっている語句を代替語句と判断するようにしてもよい。この場合も、上記のテーブル作成の手間が省けると共に、テーブルを記憶するために必要となる記憶容量を節減することができる。

できる。

[0020] 語句自体に付隨する背景知識についても、語句によって様々な性質のものが無数に存在するので、事前にそれらを全て洗い出し、カテゴリーに応じて分類することは困難である。請求項2、3記載の発明では、同じ構文の中で置き換え可能な語句を同じカテゴリーの語句と判断するので、あらゆる語句の背景知識を認識しカテゴリーに応じて分類することなく、おおよそ語句自体に付隨する背景知識を踏まえた上で適切な対訳を得ることができる。

[0021] また、請求項1又は請求項2記載の発明において、例えば請求項4に記載したように、対訳判断手段を、検索手段による検索によって抽出された自然文と原文との一致度を求め、求めた一致度に基づいて選択した自然文の対訳文における少なくとも翻訳対象語句の対訳を、原文中の少なくとも翻訳対象語句の対訳と判断するよう構成することが好ましい。これにより、検索手段による検索によって抽出された自然文のうち、原文に近い自然文（例えば原文中に存在する多義語と同じ意味で用いている自然文等）が、高い確率で一致度の高い自然文として選択され、選択された自然文に対応する自然な対訳文が得られ、該対訳文における、少なくとも翻訳対象語句の対訳が、原文中の翻訳対象語句の対訳と判断されることになる（原文との一致度等に応じて、対訳文における翻訳対象語句以外の語句の対訳も、原文中の該語句の対訳と判断してもよいことは言うまでもない）。

[0022] このように、請求項4記載の発明は、原文上のチャンクの区切りを認識することが技術的に非常に困難であることに鑑み、原言語の自然文を目的言語の対訳文と対応付けて複数記憶しておき、記憶している自然文の中から原言語の原文との一致度の高い自然文を選択し、選択した自然文に対応する自然な対訳文における対訳を、原文中の少なくとも翻訳対象語句の対訳と判断するので、原文上のチャンクの区切りを認識することなく、結果として、おおよそチャンク単位で原文の対訳を行ったに等しい自然な対訳を得ることができ、原文中の翻訳対象語句の適切な対訳（原文から目的言語の文として自然な対訳文が得られる可能性の高い対訳）を得ることができる。また、請求項4記載の発明では、検索手段による検索によって抽出された自然文と原文との一致度を求め、求めた一致度に基づいて自然文を選択することで適切な対訳を得るので、従来の機械翻訳における品詞判断や構文解析

等の複雑な処理を行う必要がなくなり、処理を簡略化することができる。

[0023] 請求項5記載の発明に係る対訳判断装置は、複数の単語から成る原言語の自然文を目的言語の対訳文と対応付けて複数記憶する記憶手段と、前記記憶手段に記憶されている原言語の複数の自然文の中から、原言語の原文中の翻訳対象語句が含まれている自然文を検索する検索手段と、前記原文中の各単語に対し前記原文中の翻訳対象語句との前記原文上での距離の小ささに応じた第1評価値を各々付与し、前記検索手段による検索によって抽出された自然文について、当該自然文の各単語のうち前記原文との一致単語を認識し、認識した一致単語に付与した前記第1評価値の合計値を自然文毎に演算し、演算した前記第1評価値の合計値に基づいて選択した自然文の対訳文における少なくとも前記翻訳対象語句の対訳を、前記原文中の少なくとも前記翻訳対象語句の対訳と判断する対訳判断手段と、を含んで構成されている。

[0024] 請求項5記載の発明は、請求項1、2と同様の記憶手段及び検索手段を備え、対訳判断手段は、原文中の各単語に対し原文中の翻訳対象語句との原文上での距離の小ささに応じた第1評価値を各々付与し、検索手段による検索によって抽出された自然文について、当該自然文の各単語のうち原文との一致単語を認識し、認識した一致単語に付与した第1評価値の合計値を自然文毎に演算し、演算した第1評価値の合計値に基づいて選択した自然文の対訳文における少なくとも翻訳対象語句の対訳を、原文中の少なくとも翻訳対象語句の対訳と判断するので、第1評価値の合計値に基づいて、原文上で翻訳対象語句に近い位置に存在している単語をより多く含む自然文、すなわち原文中の翻訳対象語句及びその近辺の単語から成るフレーズ（連語）を含んでいる可能性の高い自然文（原文との単語の並び順の類似度が高いと推定される自然文）を選択することができ、原言語の原文から目的言語の文として自然な対訳文が得られる可能性の高い、原文中の翻訳対象語句の適切な対訳（原文から目的言語の文として自然な対訳文が得られる可能性の高い対訳）を得ることができる。

[0025] また、請求項5記載の発明において、対訳判断手段を、例えば請求項6に記載したように、第1評価値の演算結果が同一の自然文が複数存在していた場合に、当該複数の自然文の各々について、前記自然文中の一致単語に対し前記自然文中の翻訳

対象語句との前記自然文上での距離の小ささに応じた第2評価値を各々付与し、前記一致単語に付与した第2評価値の合計値を自然文毎に演算し、演算した第2評価値の合計値に基づいて選択した自然文の対訳文における少なくとも翻訳対象語句の対訳を、原文中の少なくとも翻訳対象語句の対訳と判断するように構成することが好ましい。これにより、第1評価値の演算結果が同一の自然文が複数存在していた場合に、第2評価値の合計値に基づいて、原文との一致単語をより多く含みかつ一致単語の間に存在する不一致単語がなるべく少ない自然文、すなわち原文中の翻訳対象語句及びその近辺の単語から成るフレーズ（連語）を含んでいる可能性の高い自然文を選択することができ、原言語の原文から目的言語の文として自然な対訳文が得られる可能性の高い、原文中の翻訳対象語句の適切な対訳（原文から目的言語の文として自然な対訳文が得られる可能性の高い対訳）を得ることができる。

[0026] 請求項7記載の発明に係る対訳判断装置は、複数の単語から成る原言語の自然文を目的言語の対訳文と対応付けて複数記憶する記憶手段と、前記記憶手段に記憶されている原言語の複数の自然文の中から、原言語の原文中の翻訳対象語句が含まれている自然文を検索する検索手段と、前記検索手段による検索によって抽出された自然文に対し、前記原文上で前記翻訳対象語句よりも前側に存在している前側単語群のうち未探索でかつ前記翻訳対象語句との距離が最小の特定前側単語が、前記自然文上で前記翻訳対象語句又は前回の探索で抽出された前側単語の前側に所定単語数以内に存在しているか否かを探査し、前記特定前側単語が発見された場合に発見された前記特定前側単語と前記翻訳対象語句又は前回の探索で抽出された前側単語との距離の小ささに応じた第3評価値を付与することを、未探索の前側単語が前記原文上に存在しなくなる迄繰り返すと共に、前記原文上で前記翻訳対象語句よりも後側に存在している後側単語群のうち未探索でかつ前記翻訳対象語句との距離が最小の特定後側単語が、前記自然文上で前記翻訳対象語句又は前回の探索で抽出された後側単語の後側に所定単語数以内に存在しているか否かを探査し、前記特定後側単語が発見された場合に発見された前記特定後側単語が前記翻訳対象語句又は前回の探索で抽出された後側単語との距離の小ささに応じた第3評価値を付与することを、未探索の後側単語が前記原文上に存在しなくなる迄繰り返し、付与した第3評価値の合計値を自然文毎に演算し、演算した前記第3評価値の合計値に基づ

いて選択した自然文の対訳文における少なくとも前記翻訳対象語句の対訳を、前記原文中の少なくとも前記翻訳対象語句の対訳と判断する対訳判断手段と、を含んで構成されている。

[0027] 請求項7記載の発明は、請求項1、2と同様の記憶手段及び検索手段を備え、対訳判断手段は、検索手段による検索によって抽出された自然文に対し、原文上で翻訳対象語句よりも前側に存在している前側単語群のうち未探索でかつ翻訳対象語句との距離が最小の特定前側単語が、自然文上で翻訳対象語句又は前回の探索で抽出された前側単語の前側に所定単語数以内に存在しているか否かを探査し、特定前側単語が発見された場合に発見された特定前側単語と翻訳対象語句又は前回の探索で抽出された前側単語との距離の小ささに応じた第3評価値を付与することを、未探索の前側単語が原文上に存在しなくなる迄繰り返すと共に、原文上で翻訳対象語句よりも後側に存在している後側単語群のうちのうち未探索でかつ翻訳対象語句との距離が最小の特定後側単語が、自然文上で翻訳対象語句又は前回の探索で抽出された後側単語の後側に所定単語数以内に存在しているか否かを探査し、特定後側単語が発見された場合に発見された特定後側単語が翻訳対象語句又は前回の探索で抽出された後側単語との距離の小ささに応じた第3評価値を付与することを、未探索の後側単語が原文上に存在しなくなる迄繰り返し、付与した第3評価値の合計値を自然文毎に演算し、演算した第3評価値の合計値に基づいて選択した自然文の対訳文における少なくとも翻訳対象語句の対訳を、原文中の少なくとも翻訳対象語句の対訳と判断するので、第3評価値の合計値に基づいて、原文との単語の並び順の類似度が高く、かつ原文との一致単語の間に存在する不一致単語の数がなるべく少ない自然文、すなわち原文が不一致単語をなるべく含まない状態（塊度の高い状態）で含まれている自然文を選択することができ、原言語の原文から目的言語の文として自然な対訳文が得られる可能性の高い、原文中の翻訳対象語句の適切な対訳（原文から目的言語の文として自然な対訳文が得られる可能性の高い対訳）を得ることができる。

[0028] 請求項8記載の発明に係る対訳判断装置は、複数の単語から成る原言語の自然文を目的言語の対訳文と対応付けて複数記憶する記憶手段と、前記記憶手段に記憶されている原言語の複数の自然文の中から、原言語の原文中の翻訳対象語句が含まれ

ている自然文を検索する検索手段と、前記検索手段による検索によって抽出された自然文に対し、前記原文上で前記翻訳対象語句よりも前側に存在している第1の前側単語群のうち未探索でかつ前記翻訳対象語句との距離が最小の特定前側単語が、前記自然文上で前記翻訳対象語句又は前回の探索で抽出された前側単語の前側に所定単語数以内に存在しているか否かを探索し、前記特定前側単語が発見された場合に前記翻訳対象語句又は前回の探索で抽出された前側単語との距離に対応する前側についての第1の距離別単語数をカウントアップすることを、未探索の前側単語が前記第1の前側単語群の中に存在しなくなる迄繰り返すと共に、前記検索手段による検索によって抽出された自然文に対し、前記原文上で前記翻訳対象語句よりも後側に存在している第1の後側単語群のうち未探索でかつ前記翻訳対象語句との距離が最小の特定後側単語が、前記自然文上で前記翻訳対象語句又は前回の探索で抽出された後側単語の後側に所定単語数以内に存在しているか否かを探索し、前記特定後側単語が発見された場合に前記翻訳対象語句又は前回の探索で抽出された後側単語との距離に対応する後側についての第1の距離別単語数をカウントアップすることを、未探索の後側単語が前記第1の後側単語群の中に存在しなくなる迄繰り返し、前側についての第1の距離別単語数及び後側についての第1の距離別単語数から求めた総合評価に基づいて選択した自然文の対訳文における少なくとも前記翻訳対象語句の対訳を、前記原文中の少なくとも前記翻訳対象語句の対訳と判断する対訳判断手段と、を含んで構成されている。

[0029] 請求項8記載の発明は、請求項1、2と同様の記憶手段及び検索手段を備え、対訳判断手段は、検索手段による検索によって抽出された自然文に対し、原文上で翻訳対象語句よりも前側に存在している第1の前側単語群のうち未探索でかつ翻訳対象語句との距離が最小の特定前側単語が、自然文上で翻訳対象語句又は前回の探索で抽出された前側単語の前側に所定単語数以内に存在しているか否かを探索し、特定前側単語が発見された場合に翻訳対象語句又は前回の探索で抽出された前側単語との距離に対応する前側についての第1の距離別単語数をカウントアップすることを、未探索の前側単語が第1の前側単語群の中に存在しなくなる迄繰り返すと共に、検索手段による検索によって抽出された自然文に対し、原文上で翻訳対象語句よりも後側に存在している第1の後側単語群のうち未探索でかつ翻訳対象語句との距

離が最小の特定後側単語が、自然文上で翻訳対象語句又は前回の探索で抽出された後側単語の後側に所定単語数以内に存在しているか否かを探索し、特定後側単語が発見された場合に翻訳対象語句又は前回の探索で抽出された後側単語との距離に対応する後側についての第1の距離別単語数をカウントアップすることを、未探索の後側単語が第1の後側単語群の中に存在しなくなる迄繰り返し、前側についての第1の距離別単語数及び後側についての第1の距離別単語数から求めた総合評価に基づいて選択した自然文の対訳文における少なくとも翻訳対象語句の対訳を、原文中の少なくとも翻訳対象語句の対訳と判断するので、前側についての第1の距離別単語数及び後側についての第1の距離別単語数から求めた総合評価に基づいて、原文との単語の並び順の類似度が高く、かつ原文との一致単語の間に存在する不一致単語の数がなるべく少ない自然文、すなわち原文が不一致単語をなるべく含まない状態（塊度の高い状態）で含まれている自然文を選択することができ、原言語の原文から目的言語の文として自然な対訳文が得られる可能性の高い、原文中の翻訳対象語句の適切な対訳（原文から目的言語の文として自然な対訳文が得られる可能性の高い対訳）を得ることができる。

[0030] また、請求項8記載の発明において、対訳判断手段は、例えば請求項9に記載したように、検索手段による検索によって抽出された自然文に対し、原文上で翻訳対象語句よりも前側に存在しておりかつ翻訳対象語句と隣り合っている前側単語を除外した第2の前側単語群のうち未探索でかつ翻訳対象語句との距離が最小の特定前側単語が、自然文上で翻訳対象語句又は前回の探索で抽出された前側単語の前側に所定単語数以内に存在しているか否かを探索し、特定前側単語が発見された場合に発見された特定前側単語と翻訳対象語句又は前回の探索で抽出された前側単語との距離に1又は前回の探索で発見された前側単語について求めた距離を加算した距離に対応する前側についての第2の距離別単語数をカウントアップすることを、未探索の前側単語が原文上に存在しなくなる迄繰り返し、前側についての第1の距離別単語数と前側についての第2の距離別単語数のうち単語間距離がより短いとの結果を示している方を最終評価として選択すると共に、検索手段による検索によって抽出された自然文に対し、原文上で翻訳対象語句よりも後側に存在しておりかつ翻訳対象語句と隣り合っている後側単語を除外した第2の後側単語群のうち未探索でか

つ翻訳対象語句との距離が最小の特定後側単語が、自然文上で翻訳対象語句又は前回の探索で抽出された後側単語の後側に所定単語数以内に存在しているか否かを探索し、特定後側単語が発見された場合に発見された特定後側単語と翻訳対象語句又は前回の探索で抽出された後側単語との距離に1又は前回の探索で発見された後側単語について求めた距離を加算した距離に対応する後側についての第2の距離別単語数をカウントアップすることを、未探索の後側単語が原文上に存在しなくなる迄繰り返し、後側についての第1の距離別単語数と後側についての第2の距離別単語数のうち単語間距離がより短いとの結果を示している方を最終評価として選択し、前側及び後側についての最終評価を加算集計して総合評価を求め、求めた総合評価において単語間距離がより短いとの結果を示している自然文の対訳文における少なくとも翻訳対象語句の対訳を、原文中の少なくとも翻訳対象語句の対訳と判断するように構成することが好ましい。上記のように、第2の距離別単語数を第1の距離別単語数と併用することで、原文に対して一部の単語の順序が入れ替わっている自然文についても第2の距離別単語数によって原文との単語の並び順の類似度を正当に評価することができ、原文中の翻訳対象語句の適切な対訳が得られる確率が向上する。

[0031] 請求項10記載の発明に係る対訳判断方法は、目的言語の対訳文と対応付けて記憶手段に複数記憶されている、各々複数の単語から成る原言語の自然文の中から、原言語の原文中の翻訳対象語句が含まれている自然文を検索する第1のステップ、前記第1のステップにおける検索によって抽出された自然文に基づいて、前記翻訳対象語句と原言語の同一の文中に出現する頻度の高い頻出語句を認識する第2のステップ、前記第1のステップにおける検索によって抽出された自然文のうち、前記第2のステップで認識されかつ前記原文中に存在する特定の頻出語句と、前記翻訳対象語句が各々含まれている自然文の対訳文を参照することで、前記翻訳対象語句及び前記特定の頻出語句が各々含まれる自然文の対訳文における翻訳対象語句についての高頻度の対訳を認識し、認識した高頻度の対訳を、前記原文中の前記翻訳対象語句の対訳と判断する第3のステップを含んでるので、請求項1記載の発明と同様に、原文中の翻訳対象語句の適切な対訳（原文から目的言語の文として自然な対訳文が得られる可能性の高い対訳）を得ることができる。

[0032] 請求項1 1記載の発明に係る対訳判断方法は、目的言語の対訳文と対応付けて記憶手段に複数記憶されている、各々複数の単語から成る原言語の自然文の中から、原言語の原文中の翻訳対象語句が含まれている自然文を検索する第1のステップ、前記原文中に存在しかつ前記第1のステップにおける検索によって抽出された自然文中に含まれていない着目語句と代替可能な代替語句を判断する第2のステップ、前記第1のステップにおける検索によって抽出された自然文のうち、前記第2のステップで判断した代替語句及び前記翻訳対象語句が各々含まれている自然文の対訳文における少なくとも前記翻訳対象語句の対訳を、前記原文中の少なくとも前記翻訳対象語句の対訳と判断する第3のステップを含んでいるので、請求項2記載の発明と同様に、原文中の翻訳対象語句の適切な対訳（原文から目的言語の文として自然な対訳文が得られる可能性の高い対訳）を得ることができる。

[0033] 請求項1 2記載の発明に係る対訳判断方法は、目的言語の対訳文と対応付けて記憶手段に複数記憶されている、各々複数の単語から成る原言語の自然文の中から、原言語の原文中の翻訳対象語句が含まれている自然文を検索する第1のステップ、前記原文中の各単語に対し前記原文中の翻訳対象語句との前記原文上での距離の小ささに応じた第1評価値を各々付与し、前記第1のステップにおける検索によって抽出された自然文について、当該自然文の各単語のうち前記原文との一致単語を認識し、認識した一致単語に付与した前記第1評価値の合計値を自然文毎に演算し、演算した前記第1評価値の合計値に基づいて選択した自然文の対訳文における少なくとも前記翻訳対象語句の対訳を、前記原文中の少なくとも前記翻訳対象語句の対訳と判断する第2のステップを含んでいるので、請求項5記載の発明と同様に、原文中の翻訳対象語句の適切な対訳（原文から目的言語の文として自然な対訳文が得られる可能性の高い対訳）を得ることができる。

[0034] 請求項1 3記載の発明に係る対訳判断方法は、目的言語の対訳文と対応付けて記憶手段に複数記憶されている、各々複数の単語から成る原言語の自然文の中から、原言語の原文中の翻訳対象語句が含まれている自然文を検索する第1のステップ、前記第1のステップにおける検索によって抽出された自然文に対し、前記原文上で前記翻訳対象語句よりも前側に存在している前側単語群のうち未探索でかつ前記翻訳対象語句との距離が最小の特定前側単語が、前記自然文上で前記翻訳対象語句又

は前回の探索で抽出された前側単語の前側に所定単語数以内に存在しているか否かを探索し、前記特定前側単語が発見された場合に発見された前記特定前側単語と前記翻訳対象語句又は前回の探索で抽出された前側単語との距離の小ささに応じた第3評価値を付与することを、未探索の前側単語が前記原文上に存在しなくなる迄繰り返すと共に、前記原文上で前記翻訳対象語句よりも後側に存在している後側単語群のうちのうち未探索でかつ前記翻訳対象語句との距離が最小の特定後側単語が、前記自然文上で前記翻訳対象語句又は前回の探索で抽出された後側単語の後側に所定単語数以内に存在しているか否かを探索し、前記特定後側単語が発見された場合に発見された前記特定後側単語が前記翻訳対象語句又は前回の探索で抽出された後側単語との距離の小ささに応じた第3評価値を付与することを、未探索の後側単語が前記原文上に存在しなくなる迄繰り返し、付与した第3評価値の合計値を自然文毎に演算し、演算した前記第3評価値の合計値に基づいて選択した自然文の対訳文における少なくとも前記翻訳対象語句の対訳を、前記原文中の少なくとも前記翻訳対象語句の対訳と判断する第2のステップを含んでいるので、請求項7記載の発明と同様に、原文中の翻訳対象語句の適切な対訳（原文から目的言語の文として自然な対訳文が得られる可能性の高い対訳）を得ることができる。

[0035] 請求項14記載の発明に係る対訳判断方法は、目的言語の対訳文と対応付けて記憶手段に複数記憶されている、各々複数の単語から成る原言語の自然文の中から、原言語の原文中の翻訳対象語句が含まれている自然文を検索する第1のステップ、前記第1のステップにおける検索によって抽出された自然文に対し、前記原文上で前記翻訳対象語句よりも前側に存在している第1の前側単語群のうち未探索でかつ前記翻訳対象語句との距離が最小の特定前側単語が、前記自然文上で前記翻訳対象語句又は前回の探索で抽出された前側単語の前側に所定単語数以内に存在しているか否かを探索し、前記特定前側単語が発見された場合に前記翻訳対象語句又は前回の探索で抽出された前側単語との距離に対応する前側についての第1の距離別単語数をカウントアップすることを、未探索の前側単語が前記第1の前側単語群の中に存在しなくなる迄繰り返すと共に、前記第1のステップにおける検索によって抽出された自然文に対し、前記原文上で前記翻訳対象語句よりも後側に存在している第1の後側単語群のうち未探索でかつ前記翻訳対象語句との距離が最小の特定後側単

語が、前記自然文上で前記翻訳対象語句又は前回の探索で抽出された後側単語の後側に所定単語数以内に存在しているか否かを探索し、前記特定後側単語が発見された場合に前記翻訳対象語句又は前回の探索で抽出された後側単語との距離に対応する後側についての第1の距離別単語数をカウントアップすることを、未探索の後側単語が前記第1の後側単語群の中に存在しなくなる迄繰り返し、前側についての第1の距離別単語数及び後側についての第1の距離別単語数に基づいて選択した自然文の対訳文における少なくとも前記翻訳対象語句の対訳を、前記原文中の少なくとも前記翻訳対象語句の対訳と判断する第2のステップを含んでいるので、請求項8記載の発明と同様に、原文中の翻訳対象語句の適切な対訳（原文から目的言語の文として自然な対訳文が得られる可能性の高い対訳）を得ることができる。

[0036] 請求項15記載の発明に係るプログラムは、複数の単語から成る原言語の自然文を目的言語の対訳文と対応付けて複数記憶する記憶手段と接続されたコンピュータを、前記記憶手段に記憶されている原言語の複数の自然文の中から、原言語の原文中の翻訳対象語句が含まれている自然文を検索する検索手段、前記検索手段による検索によって抽出された自然文に基づいて、前記翻訳対象語句と原言語の同一の文中に出現する頻度の高い頻出語句を認識する認識手段、及び、前記検索手段による検索によって抽出された自然文のうち、前記認識手段によって認識されかつ前記原文中に存在する特定の頻出語句と、前記翻訳対象語句が各々含まれている自然文の対訳文を参照することで、前記翻訳対象語句及び前記特定の頻出語句が各々含まれる自然文の対訳文における翻訳対象語句についての高頻度の対訳を認識し、認識した高頻度の対訳を、前記原文中の前記翻訳対象語句の対訳と判断する対訳判断手段として機能させる。

請求項15記載の発明に係るプログラムは、複数の単語から成る原言語の自然文を目的言語の対訳文と対応付けて複数記憶する記憶手段と接続されたコンピュータ（記憶手段を内蔵しているコンピュータであってもよいし、記憶手段と接続された別のコンピュータと通信回線を介して接続されているコンピュータであってもよい）を、上記の検索手段、認識手段及び対訳判断手段として機能させるためのプログラムであるので、上記コンピュータが請求項15記載の発明に係るプログラムを実行することで、上記コンピュータが請求項1に記載の対訳判断装置として機能する

ことになり、請求項1記載の発明と同様に、原文中の翻訳対象語句の適切な対訳（原文から目的言語の文として自然な対訳文が得られる可能性の高い対訳）を得ることができる。

[0036] 請求項16記載の発明に係るプログラムは、複数の単語から成る原言語の自然文を目的言語の対訳文と対応付けて複数記憶する記憶手段と接続されたコンピュータを、前記記憶手段に記憶されている原言語の複数の自然文の中から、原言語の原文中の翻訳対象語句が含まれている自然文を検索する検索手段、前記原文中に存在しかつ前記検索手段による検索によって抽出された自然文中に含まれていない着目語句と代替可能な代替語句を判断する判断手段、及び、前記検索手段による検索によって抽出された自然文のうち、前記判断手段によって判断された代替語句及び前記翻訳対象語句が各々含まれている自然文の対訳文における少なくとも前記翻訳対象語句の対訳を、前記原文中の少なくとも前記翻訳対象語句の対訳と判断する対訳判断手段として機能させる。

請求項16記載の発明に係るプログラムは、上記の記憶手段と接続されたコンピュータを、上記の検索手段、判断手段及び対訳判断手段として機能させるためのプログラムであるので、上記コンピュータが請求項16記載の発明に係るプログラムを実行することで、上記コンピュータが請求項2に記載の対訳判断装置として機能することになり、請求項2記載の発明と同様に、原文中の翻訳対象語句の適切な対訳（原文から目的言語の文として自然な対訳文が得られる可能性の高い対訳）を得ることができる。

[0037] 請求項17記載の発明に係るプログラムは、複数の単語から成る原言語の自然文を目的言語の対訳文と対応付けて複数記憶する記憶手段と接続されたコンピュータを、前記記憶手段に記憶されている原言語の複数の自然文の中から、原言語の原文中の翻訳対象語句が含まれている自然文を検索する検索手段、及び、前記原文中の各単語に対し前記原文中の翻訳対象語句との前記原文上での距離の小ささに応じた第1評価値を各々付与し、前記検索手段による検索によって抽出された自然文について、当該自然文の各単語のうち前記原文との一致単語を認識し、認識した一致単語に付与した前記第1評価値の合計値を自然文毎に演算し、演算した前記第1評価値の合計値に基づいて選択した自然文の対訳文における少なくとも前記翻訳対象語

句の対訳を、前記原文中の少なくとも前記翻訳対象語句の対訳と判断する対訳判断手段として機能させる。

請求項17記載の発明に係るプログラムは、上記の記憶手段と接続されたコンピュータを、上記の検索手段及び対訳判断手段として機能させるためのプログラムであるので、上記コンピュータが請求項17記載の発明に係るプログラムを実行することで、上記コンピュータが請求項5に記載の対訳判断装置として機能することになり、請求項5記載の発明と同様に、原文中の翻訳対象語句の適切な対訳（原文から目的言語の文として自然な対訳文が得られる可能性の高い対訳）を得ることができる。

[0038] 請求項18記載の発明に係るプログラムは、複数の単語から成る原言語の自然文を目的言語の対訳文と対応付けて複数記憶する記憶手段と接続されたコンピュータを、前記記憶手段に記憶されている原言語の複数の自然文の中から、原言語の原文中の翻訳対象語句が含まれている自然文を検索する検索手段、及び、前記検索手段による検索によって抽出された自然文に対し、前記原文上で前記翻訳対象語句よりも前側に存在している前側単語群のうち未探索でかつ前記翻訳対象語句との距離が最小の特定前側単語が、前記自然文上で前記翻訳対象語句又は前回の探索で抽出された前側単語の前側に所定単語数以内に存在しているか否かを探索し、前記特定前側単語が発見された場合に発見された前記特定前側単語と前記翻訳対象語句又は前回の探索で抽出された前側単語との距離の小ささに応じた第3評価値を付与することを、未探索の前側単語が前記原文上に存在しなくなる迄繰り返すと共に、前記原文上で前記翻訳対象語句よりも後側に存在している後側単語群のうちのうち未探索でかつ前記翻訳対象語句との距離が最小の特定後側単語が、前記自然文上で前記翻訳対象語句又は前回の探索で抽出された後側単語の後側に所定単語数以内に存在しているか否かを探索し、前記特定後側単語が発見された場合に発見された前記特定後側単語が前記翻訳対象語句又は前回の探索で抽出された後側単語との距離の小ささに応じた第3評価値を付与することを、未探索の後側単語が前記原文上に存在しなくなる迄繰り返し、付与した第3評価値の合計値を自然文毎に演算し、演算した前記第3評価値の合計値に基づいて選択した自然文の対訳文における少なくとも前記翻訳対象語句の対訳を、前記原文中の少なくとも前記翻訳対象語句の対訳と判断

する対訳判断手段として機能させる。

請求項18記載の発明に係るプログラムは、上記の記憶手段と接続されたコンピュータを、上記の検索手段及び対訳判断手段として機能させるためのプログラムであるので、上記コンピュータが請求項18記載の発明に係るプログラムを実行することで、上記コンピュータが請求項7に記載の対訳判断装置として機能することになり、請求項7記載の発明と同様に、原文中の翻訳対象語句の適切な対訳（原文から目的言語の文として自然な対訳文が得られる可能性の高い対訳）を得ることができる。

[0039] 請求項19記載の発明に係るプログラムは、複数の単語から成る原言語の自然文を目的言語の対訳文と対応付けて複数記憶する記憶手段と接続されたコンピュータを、前記記憶手段に記憶されている原言語の複数の自然文の中から、原言語の原文中の翻訳対象語句が含まれている自然文を検索する検索手段、及び、前記検索手段による検索によって抽出された自然文に対し、前記原文上で前記翻訳対象語句よりも前側に存在している第1の前側単語群のうち未探索でかつ前記翻訳対象語句との距離が最小の特定前側単語が、前記自然文上で前記翻訳対象語句又は前回の探索で抽出された前側単語の前側に所定単語数以内に存在しているか否かを探索し、前記特定前側単語が発見された場合に前記翻訳対象語句又は前回の探索で抽出された前側単語との距離に対応する前側についての第1の距離別単語数をカウントアップすることを、未探索の前側単語が前記第1の前側単語群の中に存在しなくなる迄繰り返すと共に、前記検索手段による検索によって抽出された自然文に対し、前記原文上で前記翻訳対象語句よりも後側に存在している第1の後側単語群のうち未探索でかつ前記翻訳対象語句との距離が最小の特定後側単語が、前記自然文上で前記翻訳対象語句又は前回の探索で抽出された後側単語の後側に所定単語数以内に存在しているか否かを探索し、前記特定後側単語が発見された場合に前記翻訳対象語句又は前回の探索で抽出された後側単語との距離に対応する後側についての第1の距離別単語数をカウントアップすることを、未探索の後側単語が前記第1の後側単語群の中に存在しなくなる迄繰り返し、前側についての第1の距離別単語数及び後側についての第1の距離別単語数に基づいて選択した自然文の対訳文における少なくとも前記翻訳対象語句の対訳を、前記原文中の少なくとも前記翻訳対象語句の対訳と判

断する対訳判断手段として機能させる。

請求項19記載の発明に係るプログラムは、上記の記憶手段と接続されたコンピュータを、上記の検索手段及び対訳判断手段として機能させるためのプログラムであるので、上記コンピュータが請求項19記載の発明に係るプログラムを実行することで、上記コンピュータが請求項8に記載の対訳判断装置として機能することになり、請求項8記載の発明と同様に、原文中の翻訳対象語句の適切な対訳（原文から目的言語の文として自然な対訳文が得られる可能性の高い対訳）を得ることができる。

発明の効果

[0040] 以上説明したように本発明は、目的言語の対訳文と対応付けて記憶手段に複数記憶されている各々複数の単語から成る原言語の自然文の中から、原言語の原文中の翻訳対象語句が含まれている自然文を検索し、該検索によって抽出された自然文に基づいて、翻訳対象語句と原言語の同一の文中に出現する頻度の高い頻出語句を認識し、原文中に存在する特定の頻出語句と翻訳対象語句が各々含まれている自然文の対訳文を参照することで、翻訳対象語句及び特定の頻出語句が各々含まれる自然文の対訳文における翻訳対象語句についての高頻度の対訳を認識し、認識した高頻度の対訳を、原文中の翻訳対象語句の対訳と判断するので、原文中の翻訳対象語句の適切な対訳を得ることができる、という優れた効果を有する。

[0041] また本発明は、目的言語の対訳文と対応付けて記憶手段に複数記憶されている各々複数の単語から成る原言語の自然文の中から、原言語の原文中の翻訳対象語句が含まれている自然文を検索し、原文中に存在しあつ前記検索で抽出された自然文中に含まれていない着目語句と代替可能な代替語句を判断し、前記検索で抽出された自然文のうち、前記判断した代替語句及び翻訳対象語句が各々含まれている自然文の対訳文における少なくとも翻訳対象語句の対訳を、原文中の少なくとも翻訳対象語句の対訳と判断するので、原文中の翻訳対象語句の適切な対訳を得ることができる、という優れた効果を有する。

[0042] また本発明は、目的言語の対訳文と対応付けて記憶手段に複数記憶されている各々複数の単語から成る原言語の自然文の中から原言語の原文中の翻訳対象語句が含まれている自然文を検索し、原文中の各単語に対し原文中の翻訳対象語句との原文

上での距離の小ささに応じた第1評価値を各々付与し、前記検索で抽出された自然文について、当該自然文の各単語のうち原文との一致単語を認識し、認識した一致単語に付与した第1評価値の合計値を自然文毎に演算し、演算した第1評価値の合計値に基づいて選択した自然文の対訳文における少なくとも翻訳対象語句の対訳を、原文中の少なくとも翻訳対象語句の対訳と判断するので、原文中の翻訳対象語句の適切な対訳を得ることができる、という優れた効果を有する。

[0043] 更に本発明は、目的言語の対訳文と対応付けて記憶手段に複数記憶されている各々複数の単語から成る原言語の自然文の中から、原言語の原文中の翻訳対象語句が含まれている自然文を検索し、該検索で抽出された自然文に対し、原文上で翻訳対象語句よりも前側に存在している前側単語群のうち未探索でかつ翻訳対象語句との距離が最小の特定前側単語が、自然文上で翻訳対象語句又は前回の探索で抽出された前側単語の前側に所定単語数以内に存在しているか否かを探索し、特定前側単語が発見された場合に発見された特定前側単語と翻訳対象語句又は前回の探索で抽出された前側単語との距離の小ささに応じた第3評価値を付与することを、未探索の前側単語が原文上に存在しなくなる迄繰り返すと共に、原文上で翻訳対象語句よりも後側に存在している後側単語群のうちのうち未探索でかつ翻訳対象語句との距離が最小の特定後側単語が、自然文上で翻訳対象語句又は前回の探索で抽出された後側単語の後側に所定単語数以内に存在しているか否かを探索し、特定後側単語が発見された場合に発見された特定後側単語が翻訳対象語句又は前回の探索で抽出された後側単語との距離の小ささに応じた第3評価値を付与することを、未探索の後側単語が原文上に存在しなくなる迄繰り返し、付与した第3評価値の合計値を自然文毎に演算し、演算した第3評価値の合計値に基づいて選択した自然文の対訳文における少なくとも翻訳対象語句の対訳を、原文中の少なくとも翻訳対象語句の対訳と判断するので、原文中の翻訳対象語句の適切な対訳を得ることができる、という優れた効果を有する。

[0044] また本発明は、目的言語の対訳文と対応付けて記憶手段に複数記憶されている各々複数の単語から成る原言語の自然文の中から、原言語の原文中の翻訳対象語句が含まれている自然文を検索し、該検索で抽出された自然文に対し、原文上で翻訳対象語句よりも前側に存在している第1の前側単語群のうち未探索でかつ翻訳対象語

句との距離が最小の特定前側単語が、自然文上で翻訳対象語句又は前回の探索で抽出された前側単語の前側に所定単語数以内に存在しているか否かを探索し、特定前側単語が発見された場合に翻訳対象語句又は前回の探索で抽出された前側単語との距離に対応する前側についての第1の距離別単語数をカウントアップすることを、未探索の前側単語が第1の前側単語群の中に存在しなくなる迄繰り返すと共に、前記検索で抽出された自然文に対し、原文上で翻訳対象語句よりも後側に存在している第1の後側単語群のうち未探索でかつ翻訳対象語句との距離が最小の特定後側単語が、自然文上で翻訳対象語句又は前回の探索で抽出された後側単語の後側に所定単語数以内に存在しているか否かを探索し、特定後側単語が発見された場合に翻訳対象語句又は前回の探索で抽出された後側単語との距離に対応する後側についての第1の距離別単語数をカウントアップすることを、未探索の後側単語が第1の後側単語群の中に存在しなくなる迄繰り返し、前側についての第1の距離別単語数及び後側についての第1の距離別単語数から求めた総合評価に基づいて選択した自然文の対訳文における少なくとも翻訳対象語句の対訳を、原文中の少なくとも翻訳対象語句の対訳と判断するので、

原文中の翻訳対象語句の適切な対訳を得ることができる、という優れた効果を有する。

図面の簡単な説明

[0045] [図1] 本実施形態に係るPCの概略構成を示すブロック図である。

[図2] 対訳判断処理の内容を示すフローチャートである。

[図3] 対訳DBに登録されている"operation"を含む自然文と対訳文の一例を示す図表である。

[図4] ネットワークを介してクライアントPCとサーバ・コンピュータが接続されたコンピュータ・システムに本発明を適用した態様を説明するためのブロック図である。

[図5] 本発明を適用した機械翻訳装置の概略構成の一例を示すブロック図である。

発明を実施するための最良の形態

[0046] 以下、図面を参照して本発明の実施形態の一例を詳細に説明する。なお、以下では、まずユーザから翻訳対象語句が指定されると、指定された翻訳対象語句の対訳をユーザに提示する電子辞書装置に本発明を適用した態様について説明する。

[0047] 図1には、上記の電子辞書装置として機能することが可能なパーソナル・コンピュータ(PC)10が示されている。PC10はCPU10A、ROM10B、RAM10C及び入出力ポート10Dを備えており、これらがデータバス、アドレスバス、制御バス等から成るバス10Eを介して互いに接続されて構成されている。また入出力ポート10Dには、各種の入出力機器として、CRTやLCD等から成るディスプレイ12、ユーザがデータ等を入力するためキーボード14、マウス16、ハードディスクドライブ(HDD)18、CD-ROM24からのデータの読み出しを行うCD-ROMドライブ20、及び紙原稿等の読み取りが可能なスキャナ22が各々接続されている。

[0048] PC10には、PC10を電子辞書装置として機能させるための対訳判断プログラム(請求項15及び請求項16に記載のプログラムに相当)がHDD18にインストールされており、このHDD18には、前記対訳判断プログラムが対訳判断に用いるためのデータが記憶された対訳データベース(対訳DB)も記憶されている。対訳判断プログラムをPC10にインストール(移入)するには幾つかの方法が

あるが、例えば対訳判断プログラムをセットアッププログラムと共にCD-ROM
24に記録しておき、該CD-ROM

：光学的手法による文字認識）処理を経て得られたテキスト等）を適用することができる。

[0052] また、上記の所定の操作としては、例えばディスプレイ12に原文が表示されている状態で、翻訳対象語句を選択することで翻訳対象語句を反転表示させた後に、翻訳対象語句を右クリックすることで表示されるコンテキストメニュー内の「対訳出力」に相当する項目を選択する等の操作を適用することができる。また、上記のようにコンテキストメニューを利用することに代えて、例えばツールバー内に表示されている所定のアイコンを選択する等の操作を行うことで、翻訳対象対象語句の対訳の出力を指示可能とすることも可能である。翻訳対象語句の対訳の出力を指示する上記のような操作が行われると、PC10のCPU10Aによって対訳判断プログラムが実行されることで、図2に示す対訳判断処理が行われる。なお、この対訳判断処理は請求項10及び請求項11に記載の対訳判断方法が適用された処理であり、この処理を行うことでPC10は電子辞書装置（請求項1～請求項4に記載の対訳判断装置）として機能する。

[0053] 対訳判断処理では、まずステップ100において、指定された翻訳対象語句を含む単一の原文（処理対象の原文）のテキストデータを取り込むと共に、取り込んだ処理対象の原文中の翻訳対象語句を識別する情報を取り込む。なお、処理対象の原文は、翻訳対象語句を含む文であってもよいし、翻訳対象語句を含む文節、連語、定型表現、コロケーションの何れであってもよいが、文節、連語、定型表現、コロケーションの何れかを処理対象の原文とする場合、処理対象の原文としての文節、連語、定型表現、コロケーションは、ユーザによって指定させるようにしてもよいし、対訳判断処理で自動的に判断することも可能である。

[0054] 次のステップ102では、ステップ100で取り込んだ処理対象の原文のテキストデータをキーにして、対訳DBに登録されている自然文の中から処理対象の原文と完全一致している自然文を検索すると共に、翻訳対象語句のテキストデータをキーにして、対訳DBに登録されている自然文の中から翻訳対象語句が含まれている自然文（少なくとも翻訳対象語句が含まれていることで原文と部分一致している自然文）を検索する。また、翻訳対象語句が複数の単語から構成されている場合には、翻訳対象

語句を構成する複数の単語のうちの少なくとも1つが含まれている自然文も同時に検索される。なお、ステップ102は、請求項1, 2, 15, 16に記載の検索手段に対応しており、請求項10, 11に記載の第1のステップにも対応している。次のステップ104～ステップ112は請求項4に記載の対訳判断手段に対応している。

[0055] ステップ104では、ステップ102の検索により処理対象の原文と完全一致している自然文が対訳DBから抽出されたか否か判定する。この判定が肯定された場合はステップ106へ移行し、処理対象の原文と完全一致している自然文と対応付けられて対訳DBに登録されている目的言語の対訳文を読み出してディスプレイ12に表示し、処理を終了する。なお、この対訳文の表示に際しては、読み出した対訳文上での翻訳対象語句の対訳が認識され、認識された翻訳対象語句の対訳が強調表示される。この場合、ユーザは指定した翻訳対象語句の適切な対訳（原文から目的言語の文として自然な対訳文が得られる可能性の高い対訳）を認識できると共に、指定した翻訳対象語句を含む処理対象の原文の適切な対訳文（目的言語の文として自然な対訳文）を認識することができる。

[0056] (例1) 上記処理について、実例を挙げて更に説明する。原言語が英語、目的言語が日本語であり、翻訳対象語句として、

"For your safety, don't rush into the train."

という処理対象の原文に含まれる"For your safety"が指定された場合、通常の翻訳プロセスでは、原言語と対訳を単語単位で登録しているので、上記の翻訳対象語句を構成する単語"safety"に対しては対訳「安全」が選択される。しかし、本実施形態に係る対訳判断処理では、対訳DBに原言語の自然文と目的言語の対訳文を対応付けて登録しているので、上記の原文と完全一致している自然文が対訳DBに登録されている可能性がある。そして、完全一致している自然文が対訳DBに登録されていた場合、当該自然文と対応付けられて対訳DBに登録されている対訳文が、以下のように翻訳対象語句の対訳が強調表示されて出力される。

「危険ですから、駆け込み乗車はおやめ下さい。」

通常の翻訳プロセスでは、単語 "safety" の対訳として「危険」を選択することは困難であり、「あなたの安全のため、その列車に駆け込まないでください。」等の不自然な対訳文しか得られないが、本発明によれば母語として自然な対訳文が得られる。

[0057] なお、ステップ106における対訳文の表示において、処理対象の原文と完全一致している自然文に加えて、ステップ102の検索によって抽出された処理対象の原文と部分一致している自然文も併せて表示するようにしてもよい。処理対象の原文と部分一致している自然文は検索によって通常複数抽出されるが、これらを表示する場合には、処理対象の原文の部分一致している個々の自然文に対して次に述べる一致度を各々演算し、各自然文の対訳文を一致度の降順にディスプレイ12に表示することが好ましい。

[0058] 一方、ステップ102の検索により処理対象の原文と完全一致している自然文が対訳DBから抽出されなかった場合には、ステップ104の判定が否定されてステップ108へ移行し、ステップ102の検索により対訳DBから抽出された各自然文（処理対象の原文と部分一致している各自然文）について、自然文の各単語のうち原文中の何れかの単語と一致する単語の数（一致単語数）を各々計数し、一致単語数の計数結果に基づいて原文との一致度を各々演算する。

[0059] なお、本実施形態では、対訳判断プログラムのインストール時に、原言語の自然文中に頻出する単語（例えば英語の自然文における "a", "the", "to", "in" 等）が登録された頻出単語テーブルもHDD18に記憶される。そして、ステップ108における一致単語数の計数は頻出単語テーブルを参照しながら行われ、頻出単語テーブルに登録されている単語は一致単語数の計数対象から除外される。これにより、頻出単語が一致単語数に及ぼす影響を排除することができる。

[0060] また、本実施形態では、対訳判断プログラムのインストール時に、単複又は時制の相違により語尾が相違している単語が登録された語尾変化単語テーブルもHDD18に記憶される。そして、ステップ108における一致単語数の計数では、語尾のみが不一致の単語が出現した場合に、語尾変化単語テーブルが参照されることで、語尾の不一致が単複又は時制の相違によるものか否かが判断され、単複又は時制の相違

により語尾が相違している単語は一致単語とみなして計数される。なお、非定型の語尾変化のみ語尾変化単語テーブルに登録し、定型の語尾変化（例えば英単語における複数形の"s"や過去形の"ed"等）は自動的に一致単語と判断するようにしてもよい。これにより、単複又は時制の相違により語尾が相違している単語（本来は一致とみなすべき単語）が一致単語数に及ぼす影響も排除することができる。

- [0061] また、本実施形態では、ステップ108における一致単語数の計数に際し、原文中の何れかの単語と一致したことで一致単語数に一旦計上した単語は、自然文中に再度出現したとしても一致単語数に計上しないことにより、複数回出現した一致単語を重複計数しないようにしている。これにより、同一の一致単語が自然文中の複数箇所に存在していた場合にも、この一致単語が一致単語数に及ぼす影響を排除することができる。
- [0062] また、一致度の演算式は、一致単語数が多くなるに従って一致度が高くなるように定めることができ、例えば一致単語数を、翻訳対象語句を構成する単語数で正規化する演算式（一致度=一致単語数÷翻訳対象語句を構成する単語数）を用いることができる。
- [0063] 次のステップ110では、各自然式について演算した一致度を比較することで、一致度が最大の自然文が複数存在しているか否か判定する。判定が否定された場合にはステップ112へ移行し、一致度最大の自然文と対応付けられて対訳DBに登録されている目的言語の対訳文を読み出し、読み出した対訳文上での翻訳対象語句の対訳を認識し、認識された翻訳対象語句の対訳が強調表示されるように、読み出した対訳文をディスプレイ12に表示して処理を終了する。この場合、ユーザは指定した翻訳対象語句の適切な対訳（原文から目的言語の文として自然な対訳文が得られる可能性の高い対訳）を認識することができる。
- [0064] なお、ステップ112では、ステップ102の検索によって抽出された自然文のうち、一致度が最大の自然文以外に、一致度の降順に複数の自然文（一致度の降順に一定

国家」の場合とは異なる以下の順序で表示される。

1. 「発」 (不一致単語数「0」)
2. 「発展」 (不一致単語数「1」)
3. 「発展中」 (不一致単語数「2」)
4. 「発展中国」 (不一致単語数「3」)
5. 「発展中国家」 (不一致単語数「4」)

同様に、原文が「発展」で翻訳対象語句が「発展」の場合にも、一致度及び不一致単語数に基づき、原文が「発展中国家」や「発」の場合とは異なる以下の順序で表示される。

1. 「発展」 (一致単語数「2」、一致度100%、不一致単語数「0」)
2. 「発展中」 (一致単語数「2」、一致度100%、不一致単語数「1」)
3. 「発展中国」 (一致単語数「2」、一致度100%、不一致単語数「2」)
4. 「発展中国家」 (一致単語数「2」、一致度100%、不一致単語数「3」)
5. 「発」 (一致単語数「1」、一致度50%、不一致単語数「0」)

このように、翻訳対象語句が同一であっても原文が相違していれば原文に応じた適切な順序で対訳文が一覧表示されるので、ユーザは対訳文の一覧表示結果を参照することで、原文に応じた適切な対訳を認識することができる。

[0069] また、一致度が最大の自然文が複数存在していた場合、ステップ110の判定が肯定されてステップ114へ移行し、ステップ114, 116において、翻訳対象語句と原言語の同一の文中に出現する頻度の高い単語（翻訳対象語句とのコリレーションの高い単語：請求項1, 10, 15に記載の頻出語句に相当）に基づいて自然文を選択するコリレーション解析を行う。

[0070] すなわち、ステップ114では、ステップ102の検索により対訳DBから抽出された各自然文に対し、処理対象の原文中の翻訳対象語句以外の各単語（但し、頻出単語テーブルに登録されている頻出単語は除外）が含まれているか否かを検索することで、翻訳対象語句と原言語の同一の文中に出現する頻度の高い単語（翻訳対象語句とのコリレーションの高い単語）が処理対象の原文に含まれているか否かを探索する。この処理では、処理対象の原文中の翻訳対象語句以外の各単語のうち、例えば

各自然文中における出現回数が所定回（1回でも複数回でもよい）以上の単語を、翻訳対象語句とのコリレーションの高い単語と判断することができる。なお、上記のステップ114は請求項1, 15に記載の認識手段に対応しており、請求項10に記載の第2のステップにも対応している。

[0071] 次のステップ116では、ステップ114における探索により翻訳対象語句とのコリレーションの高い単語が見つかったか否か判定する。判定が肯定された場合はステップ118へ移行し、ステップ102の検索により対訳DBから抽出された各自然文のうち、翻訳対象語句及びステップ114で判断した翻訳対象語句とのコリレーションが高い単語（処理対象の原文中に存在している単語）が各々含まれている自然文について、その対訳文を参照し、上記の自然文における翻訳対象語句についての高頻度の対訳を判断する。なお、翻訳対象語句についての高頻度の対訳は、具体的には、翻訳対象語句及び翻訳対象語句とのコリレーションが高い単語が各々含まれている自然文として单一の自然文が抽出された場合には、該自然文における翻訳対象語句の対訳を高頻度の対訳と判断し、上記の自然文として複数の自然文が抽出された場合には、これらの自然文における翻訳対象語句の対訳のうち最も出現頻度の高い対訳を高頻度の対訳と判断することができる。

[0072] 上記のように、翻訳対象語句が含まれており、かつ翻訳対象語句とのコリレーションが高い単語として処理対象の原文と同一の単語が含まれている自然文が存在している場合、当該自然文は翻訳対象語句を処理対象の原文と同じ意味で用いている文である可能性が高いが、上記の自然文の中に、翻訳対象語句についての対訳が相違している自然文が混在している可能性もある。これに対してステップ118では、翻訳対象語句及び翻訳対象語句とのコリレーションが高い単語が各々含まれている自然文の対訳文を参照して、上記の自然文における翻訳対象語句についての高頻度の対訳を判断するので、処理対象の原文における翻訳対象語句の適切な対訳を得ることができる。

[0073] そしてステップ119では、翻訳対象語句及び翻訳対象語句とのコリレーションが高い単語が各々含まれている自然文のうち、翻訳対象語句がステップ118で認識した高頻度の対訳と対応付けられている自然文と対応付けられて対訳DBに登録されて

いる目的言語の対訳文を読み出し、読み出した対訳文上での翻訳対象語句及び翻訳対象語句とのコリレーションが高い単語の対訳が強調表示されるように、読み出した対訳文をディスプレイ12に表示して処理を終了する。この場合もユーザは指定した翻訳対象語句の適切な対訳（原文から目的言語の文として自然な対訳文が得られる可能性の高い対訳）を認識することができる。なお、上記のステップ116～ステップ119は、請求項1、15に記載の対訳判断手段、請求項10に記載の第3のステップに各々対応している。

[0074] なお、ステップ119においても、ステップ102の検索によって抽出された自然文のうち、翻訳対象語句及び翻訳対象語句とのコリレーションが高い単語が各々含まれており、翻訳対象語句が認識した高頻度の対訳と対応付けられている自然文以外に、一致度の降順に複数の自然文（一致度の降順に一定数の自然文、或いは一致度が所定値以上の全ての自然文）の対訳文も対訳DBから読み出し、一致度及び不一致単語数に応じた順序でディスプレイ12に併せて一覧表示する。

[0075] （例5）上記処理について実例を挙げて更に説明する。原言語が英語、目的言語が日本語であり、翻訳対象語句として、

"I had an operation to remove a rectal cancer."

という処理対象の原文に含まれる"operation"が指定されると共に、翻訳対象語句"operation"を含む自然文として図3に示す自然文(1)～(8)及び他の自然文が各々対訳文と共に対訳DBに登録されていた場合、自然文(6), (8)は何れも原文との一致単語数が「2」であるので、一致度のみから最適な自然文（対訳文）を選択することは困難である。ここで、処理対象の原文中の翻訳対象語句"operation"以外の各単語（例えば"remove", "rectal", "cancer"等）をキーにして翻訳対象語句"operation"を含む自然文(1)～(8)を検索すると、自然文(6), (8)に"cancer"が含まれていることから、処理対象の原文における翻訳対象語句"operation"とのコリレーションの高い単語として"cancer"が抽出される。

[0076] ここで、翻訳対象語句"operation"及びコリレーションが高い単語"cancer"が含まれている自然文は翻訳対象語句"operation"を処理対象の原文と同じ意味で用いている文である可能性が高いが、対訳DBに登録されている上記の条件を満たす自然文

容に応じて"low" (法) 等の単語が抽出されることになる。

[0079] 一方、翻訳対象語句とのコリレーションの高い単語が抽出されなかつた場合には、ステップ116の判定が否定されてステップ120へ移行し、ステップ120～134において、処理対象の原文中の翻訳対象語句以外の着目単語と代替可能な代替単語（請求項2, 3, 11, 16に記載の代替語句に相当）に基づいて自然文を選択するスキーム解析を行う。

[0080] すなわち、ステップ120では処理対象の原文中に存在し、かつステップ102の検索により対訳DBから抽出された各自然文中に存在しない単語（着目単語と称する）を判断する。ステップ122では、対訳DBに登録されている自然文の中から着目単語が含まれている自然文を検索する。なお、着目単語が複数になる場合もあるが、この場合は各着目単語についてステップ122の検索を行う。またステップ124では、ステップ122の検索により抽出された各自然文中の着目単語のみが相違している自然文（請求項3に記載の「着目語句を含む自然文の検索によって抽出された自然文と同一の構文の自然文」に相当）を各々検索する検索条件を設定し、設定した検索条件に基づき対訳DBに登録されている自然文の中から該当する自然文の検索を行う。

[0081] 次のステップ126では、ステップ124の検索によって該当する自然文が抽出されたか否か判定する。着目単語が別の単語に入れ替わっている以外は同一の自然文が対訳DBに各々存在存在していた場合、前記別の単語は着目単語に代えて使用可能な代替単語である可能性が高い。このため、判定が肯定された場合はステップ128へ移行し、検索により抽出された自然文において、着目単語と置き換わっている単語を着目単語の代替単語と認識する。なお、上述したステップ120～130は請求項2（詳しくは請求項3）及び請求項16に記載の判断手段、請求項11に記載の第2のステップに対応している。

[0082] そしてステップ130では、一致度最大（これに代えて「一致度所定値以上」としてもよい）の複数の自然文の中に着目単語が代替単語に置き換わっている自然文が有るか否か判定する。一致度最大の複数の自然文の中に、処理対象の原文中に存在する着目単語に代えて代替単語が含まれている自然文が存在していた場合、当該自然文は翻訳対象語句を処理対象の原文と同じ意味で用いている文である可能性が

高いと判断できる。このため、判定が肯定された場合はステップ132へ移行し、一致

度最大（これに代えて「一致度所定値以上」としてもよい）でかつ着目単語に代えて代替単語が含まれている自然文と対応付けられて対訳DBに登録されている目的言語の対訳文を読み出し、読み出した対訳文上での翻訳対象語句及び代替単語の対訳を認識し、認識された翻訳対象語句の対訳が強調表示されると共に、代替単語の対訳も代替単語であることが識別可能にマーキングされて表示されるように、読み出した対訳文をディスプレイ12に表示して処理を終了する。この場合もユーザは指定した翻訳対象語句の適切な対訳（原文から目的言語の文として自然な対訳文が得られる可能性の高い対訳）を認識することができる。なお、上述したステップ130、132は請求項2、16に記載の対訳判断手段、請求項11に記載の第3のステップに対応している。

[0083] なお、ステップ132においても、ステップ102の検索によって抽出された自然文のうち、一致度最大でかつ着目単語に代えて代替単語が含まれている自然文以外に、一致度の降順に複数の自然文（一致度の降順に一定数の自然文、或いは一致度が所定値以上の全ての自然文）の対訳文も対訳DBから読み出し、一致度及び不一致単語数に応じた順序でディスプレイ12に併せて一覧表示する。

[0084] （例6）上記処理について実例を挙げて更に説明する。原言語が英語、目的言語が日本語であり、“have lunch”を含む処理対象の原文における“have”が翻訳対象語句として指定されたが、対訳DBには“have breakfast”を含む自然文（翻訳対象語句“have”を処理対象の原文と同じ意味で用いている自然文）は登録されているものの、“have lunch”を含む自然文は登録されておらず、一致度に基づく判定やコリレーション解析によっても適正な対訳文を特定できなかった場合には、処理対象の原文における翻訳対象語句（“have”）や翻訳対象語句とのコリレーションが高いと判断した単語以外の各単語（例えば“lunch”等）が着目単語とされ、各着目単語が含まれている自然文が検索される。これにより、例えば着目単語“lunch”については、例えば“eat lunch”や“take a late lunch”を含む自然文が抽出される。

[0085] 次に、前記検索によって抽出された各自然文について、個々の自然文中の着目単語のみが相違している自然文が各々検索される。これにより、例えば着目単語“lunch”を含む自然文として抽出された自然文については、着目単語のみが相違している自然文として、例えば“eat breakfast”や“take a late breakfast”を含む自

然文

て単語「4」が着目単語「3」と代替可能である可能性の高い代替単語であると判断されることで、最初の検索（翻訳対象語句「かける」を含む自然文の検索）で抽出された自然文のうち、着目単語「3」が代替単語「4」に入れ替わっている自然文「4をかける」と対応付けられて対訳DBに登録されている対訳文が、翻訳対象語句「かける」の対訳（"multiply"）が強調表示されると共に、代替単語「4」の対訳（"four"）がマーキングされて出力されることになる。

"multiply by *four*"

従って、この例でも翻訳対象語句「かける」を処理対象の原文と同じ意味で用いている適切な自然文（対訳文）が選択される。

[0090] 一方、着目単語のみが相違している自然文の検索において該当する自然文が抽出されなかった場合（ステップ126の判定が否定された場合）、又は、一致度最大の複数の自然文の中に着目単語が代替単語に置き換わっている自然文が存在していなかった場合（ステップ130の判定が否定された場合）には、最適な対訳文を特定することは困難であるため、ステップ134へ移行し、ステップ102の検索によって抽出された全ての自然文、或いは一致度の降順に複数の自然文（一致度の降順に一定数の自然文、或いは一致度が所定値以上の全ての自然文）と対応付けられて対訳DBに登録されている対訳文を各々読み出し、読み出した個々の対訳文上の翻訳対象語句の対訳を認識し、認識した翻訳対象語句の対訳を強調表示させた状態で、個々の対訳文を一致度及び不一致単語数に応じた順序で一致度最大の各自然文一致度の降順にディスプレイ12に一覧表示して処理を終了する。この場合、ユーザは指定した翻訳対象語句についての幾つかの適切な対訳候補を認識することができる。

[0091] なお、上記では一致単語数のみを用いて一致度を演算する様態を例に説明したが、これに限定されるものではなく、自然文と原文との一致単語数が多くなるに従って一致度が高くなり、かつ自然文と原文との不一致単語数が少なくなるに従って一致度が高くなるように一致度の演算式を定め、一致単語数及び不一致単語数に応じて一致度を演算・評価するようにしてもよい。また、当初は一致単語数のみに基づいて一致度の評価を行い、コリレーション

ン解析やスキーム解析を行っても单一の自然文（対訳文）を選択することが困難な場合に不一致単語数を計数し、原文との不一致単語数が最小の自然文（の対訳文）を選択するようにしてもよい。

[0092] また、一致度の評価に際し、上記の一致単語数や不一致単語数以外に、自然文と原文との単語の並び順の類似度、或いは原文との一致単語の間に存在する自然文中の不一致単語数も併用し、自然文と原文との単語の並び順の類似度が高くなるに従って原文との一致度が高くなるように一致度を評価したり、一致単語の間に存在する不一致単語数が少なくなるに従って原文との一致度が高くなるように一致度を評価するようにしてもよい。自然文と原文との単語の並び順の類似度、及び原文との一致単語の間に存在する自然文中の不一致単語数も考慮した一致度の評価は、具体的には、例えば以下の処理によって実現することができる。

[0093] すなわち、まず原文中の各単語に、原文上での翻訳対象語句との隔たり（単語数）に応じて第1評価値を付与する。この第1評価値は、原文中の翻訳対象語句に対する第1評価値が最大となり、原文中の翻訳対象語句以外の各単語の第1評価値については、原文上での翻訳対象語句との隔たりが大きくなるに従って（間に存在している単語数が多くなるに従って）値が小さくなるように定めることができる。例えば原文が単語A～単語Jの9個の単語から成り（原文=（A, B, C, D, E, F, G, H, J））、原文中の4番目の単語Dが翻訳対象語句として指定された場合には、原文中の各単語A～単語Jに対して各々以下のような第1評価値を付与することができる。

$$A = 1.0, B = 2.0, C = 5.0, D = 10.0, E = 5.0, F = 2.0, G = 1.0, H = 0.5, J = 0.2$$

[0094] 次に、ステップ102の検索により対訳DBから抽出された翻訳対象語句を含む各自然文について、自然文の各単語のうち原文中の何れかの単語と一致する一致単語を認識し、認識した各一致単語に対して先に付与した第1評価値の合計値を演算する。そして、この第1評価値の合計値を一致度とし、検索により抽出された各自然文の対訳文を、対応する各自然文の一致度（第1評価値の合計値）の降順に出力する。例えば検索により以下の単語配列の自然文1～自然文5が抽出された場合（但し単

語Xは任意の単語)、

自然文1 = (D, X, X, E, F, G)

自然文2 = (D, X, E, F, G)

自然文3 = (D, E, F, G)

自然文4 = (A, B, X, C, X, D)

自然文5 = (D, E, F, G, H, J)

各自然文の第1評価値の合計値は、

自然文1 : 10.0(単語D) + 0(単語X) + 0(単語X) + 5.0(単語E) + 2.0(単語F) + 1.0(単語G) = 18.0

自然文2 : 10.0(単語D) + 0(単語X) + 5.0(単語E) + 2.0(単語F) + 1.0(単語G) = 18.0

自然文3 : 10.0(単語D) + 5.0(単語E) + 2.0(単語F) + 1.0(単語G) = 18.0

自然文4 : 1.0(単語A) + 2.0(単語B) + 0(単語X) + 5.0(単語C) + 0(単語X) + 10.0(単語D) = 18.0

自然文5 : 10.0(単語D) + 5.0(単語E) + 2.0(単語F) + 1.0(単語G) + 0.5(単語H) + 0.2(単語J) = 18.7

となり、第1評価値の合計値(一致度)は自然文5が最大となる。

[0095] 第1評価値は、原文中の翻訳対象語句以外の各単語に対し、原文上での翻訳対象語句との隔たりが小さくなるに従って値が大きくなるように定めているので、上記のように、翻訳対象語句を含む各自然文を第1評価値の合計値に基づいて評価することで、原文上で翻訳対象語句に近い位置に存在している単語をより多く含む自然文、すなわち原文中の翻訳対象語句及びその近辺の単語から成るフレーズ(連語)を含んでいる可能性の高い自然文(原文との単語の並び順の類似度が高いと推定される自然文)を、より一致度の高い自然文として評価することができる。なお、上記態様は請求項5, 12, 17の発明に対応している。

[0096] また、上記の例では自然文1～自然文4の一致度(第1評価値の合計値)が同値となっているが、このように第1評価値に基づく一致度が同値の自然文が複数存在していた場合には、自然文の各単語のうち原文中の何れかの単語と一致する一致単語に対し、各自然文上での翻訳対象語句との隔たり(単語数)に応じて第2評価値

を付与することを、第1評価値に基づく一致度が同値の各自然文について各々行う。
。この

自然文を、より一致度の高い自然文として評価することができる。なお、上記態様も請求項5（詳しくは請求項6）、12、17の発明に対応している。

[0099] また、自然文と原文との単語の並び順の類似度、及び原文との一致単語の間に存在する自然文中の不一致単語数も考慮した一致度の評価は、例えば以下の処理によって実現することも可能である。

[0100] すなわち、まず原文上で翻訳対象語句よりも前側に存在している前側単語群から翻訳対象語句との隔たりが最小の前側単語（このときは翻訳対象語句と隣り合っている前側単語）を抽出すると共に、原文上で翻訳対象語句よりも後側に存在している後側単語群から翻訳対象語句との隔たりが最小の後側単語（このときは翻訳対象語句と隣り合っている後側単語）を抽出する。そして、検索により対訳DBから抽出された翻訳対象語句を含む各自然文に対し、各自然文において翻訳対象語句が存在している位置を基準位置とし（自然文中の複数箇所に翻訳対象語句が各々存在している場合は何れかの翻訳対象語句が存在している位置を基準位置とする）、先に抽出した前側単語が基準位置から前側へ所定単語数の範囲内（例えば3単語以内）に存在しているか否かを探索すると共に、先に抽出した後側単語が基準位置から後側へ所定単語数の範囲内（例えば3単語以内）に存在しているか否かを探索する。そして、上記の探索によって前側単語及び後側単語が発見された自然文に対し、基準位置と前側単語及び後側単語の隔たりが小さくなるに従って値が大きくなるように定めた第3評価値を付与する。

[0101] 続いて、原文中の前側単語群から、未抽出でかつ翻訳対象語句との隔たりが最小の前側単語（このときは翻訳対象語句との間に单一の単語（前回の探索に用いた前側単語）が存在している前側単語）を抽出すると共に、原文中の後側単語群から、未抽出でかつ翻訳対象語句との隔たりが最小の後側単語（このときは翻訳対象語句との間に单一の単語（前回の探索に用いた後側単語）が存在している後側単語）を抽出する。そして、前回の探索によって前側単語及び後側単語が発見された各自然文に対し、先に抽出した前側単語が前回の探索で発見された前側単語の位置から前側へ所定単語数の範囲内（例えば3単語以内）に存在しているか否かを探索すると共に、先に抽出した後側単語が前回の探索で発見された後側単語の位置から後側へ所定単語数の範囲内（例えば3単語以内）に存在しているか否かを探索する。そし

在しているか否かが各々探索され、前側単語B及び後側単語Dが発見された自然文に対して第3評価値が付与される。続いて、前側単語として単語Aが、後側単語として単語Eが原文から各々抽出され、前回の探索で前側単語B及び後側単語Dが発見された自然文に対し、前側単語Aが前側単語Bから前側へ所定単語数の範囲内（例えば3単語以内）に存在しているか否か、及び、後側単語Eが後側単語Dから後側へ所定単語数の範囲内（例えば3単語以内）に存在しているか否かが各々探索され、前側単語A及び後側単語Eが発見された自然文に対して第3評価値が付与される。

[0105] 以下に示す10個の自然文a～自然文kに対して上記の処理を行って第3評価値の合計値の降順に並べたとすると、以下の順序となる。なお、以下の例において、単語Xは任意の単語、「/」は区切記号を意味している。

1. 自然文a = (A, B, C, D, E)
2. 自然文b = (A, B, C, D, /, B, C, D, E)
3. 自然文c = (B, C, D)
4. 自然文d = (A, B, C, X, D, E)
5. 自然文e = (A, B, C, X, D)
6. 自然文f = (A, B, C, /, C, D, E)
7. 自然文g = (A, X, B, C, X, D)
8. 自然文h = (B, C, /, C, D)
9. 自然文j = (B, X, C, /, C, X, D)
10. 自然文k = (B, X, X, C, /, C, X, X, D)

翻訳対象語句を含む各自然文に対し、上記のように第3評価値を付与し、第3評価値の合計値に基づいて評価することで、原文との単語の並び順の類似度が高く、かつ原文との一致単語の間に存在する不一致単語の数がなるべく少ない自然文、すなわち原文が不一致単語をなるべく含まない状態（塊度の高い状態）で含まれている自然文を、より一致度の高い自然文として評価することができる。なお、上記態様は請求項7, 13, 18の発明に対応している。

[0106] 更に、自然文と原文との単語の並び順の類似度、及び原文との一致単語の間に存在する自然文中の不一致単語数も考慮した一致度の評価は、例えば以下の処理に

評価に基づいて並べ替えを行ったとすると、以下の順序となる。

1. 自然文 $a = (A, B, C, D, E)$
2. 自然文 $b = (A, B, C, D, /, B, C, D, E)$
3. 自然文 $c = (A, B, C, X, D)$
4. 自然文 $d = (B, C, D)$
5. 自然文 $e = (A, X, B, C, X, D)$
6. 自然文 $f = (B, C, X, D)$
7. 自然文 $g = (B, C, /, C, D)$
8. 自然文 $h = (B, X, C, /, C, X, D)$
9. 自然文 $j = (B, X, X, C, /, C, X, X, D)$

なお、上記処理において、第2の距離別単語数の計数は必須ではなく、例えば原文に対して一部の単語の順序が入れ替わっている自然文を評価する必要がない等の場合には、第2の距離別単語数の計数を省略し、第1の距離別単語数をそのまま最終評価として用いてもよい。なお、上記態様は請求項8, 14, 19の発明に対応しており、特に第1の距離別単語数と第2の距離別単語数を併用する態様は請求項9記載の発明に対応している。

[0126] また、上述した第1評価値及び第2評価値／第3評価値／第4評価値（第1の距離別単語数及び第2の距離別単語数）に基づく一致度の評価・自然文の選択又は並べ替えは、コリレーション解析やスキーム解析による自然文の評価・選択又は並べ替えと併用できることは言うまでもないが、第1評価値と第2評価値又は第3評価値に基いて自然文の一致度を評価する態様は、原文と自然文との一致単語数のみを用いて一致度を評価する態様等と比較して、原文中に存在する連語やフレーズ等、原文全体よりも狭い範囲の語句について参考となる対訳文を検索する等の場合により適しており、例えば通常の検索モード以外にフレーズ検索モード等の検索モードを設け、このフレーズ検索モードが選択された場合に第1評価値と第2評価値又は第3評価値に基づく一致度の評価・自然文の選択又は並べ替えを行うようにしてもよい。

[0127] また、上記ではまず一致度を演算して自然文の選択（対訳の判断）を行い、一致度では適切な自然文を絞り込めない場合にコリレーション解析による自然文の選択

(対訳の判断)を行い、コリレーション解析でも適切な自然文を絞り込めない場合にスキーム解析による自然文の選択(対訳の判断)を行う例を説明したが、これに限定され

請求の範囲

[1] (補正後) 複数の単語から成る原言語の自然文を目的言語の対訳文と対応付けて複数記憶する記憶手段と、

前記記憶手段に記憶されている原言語の複数の自然文の中から、原言語の原文中の翻訳対象語句が含まれている自然文を検索する検索手段と、

前記検索手段による検索によって抽出された自然文に基づいて、前記翻訳対象語句と原言語の同一の文中に出現する頻度の高い頻出語句を認識する認識手段と、

前記検索手段による検索によって抽出された自然文のうち、前記認識手段によつて認識されかつ前記原文中に存在する特定の頻出語句と、前記翻訳対象語句が各々含まれている自然文の対訳文を参照することで、前記翻訳対象語句及び前記特定の頻出語句が各々含まれる自然文の対訳文における翻訳対象語句についての高頻度の対訳を認識し、認識した高頻度の対訳を、前記原文中の前記翻訳対象語句の対訳と判断する対訳判断手段と、

を含む対訳判断装置。

[2] (補正後) 複数の単語から成る原言語の自然文を目的言語の対訳文と対応付けて複数記憶する記憶手段と、

前記記憶手段に記憶されている原言語の複数の自然文の中から、原言語の原文中の翻訳対象語句が含まれている自然文を検索する検索手段と、

前記原文中に存在しかつ前記検索手段による検索によって抽出された自然文中に含まれていない着目語句と代替可能な代替語句を判断する判断手段と、

前記検索手段による検索によって抽出された自然文のうち、前記判断手段によつて判断された代替語句及び前記翻訳対象語句が各々含まれている自然文の対訳文における少なくとも前記翻訳対象語句の対訳を、前記原文中の少なくとも前記翻訳対象語句の対訳と判断する対訳判断手段と、

を含む対訳判断装置。

[3] (補正後) 前記判断手段は、前記着目語句を含む自然文を前記記憶手段に記憶されている複数の自然文の中から検索し、該検索によって抽出された自然文と同一の構文の自然文を前記記憶手段に記憶されている自然文の中から検索し、該検索によ

って抽出された自然文中で着目語句と置き換わっている語句を、前記代替語句と判断することを特徴とする請求項2記載の対訳判断装置。

[4] (補正後) 前記対訳判断手段は、前記検索手段による検索によって抽出された自然文と前記原文との一致度を求め、求めた一致度に基づいて選択した自然文の対訳文における少なくとも前記翻訳対象語句の対訳を、前記原文中の少なくとも前記翻訳対象語句の対訳と判断することを特徴とする請求項1又は請求項2記載の対訳判断装置。

[5] (補正後) 複数の単語から成る原言語の自然文を目的言語の対訳文と対応付けて複数記憶する記憶手段と、

前記記憶手段に記憶されている原言語の複数の自然文の中から、原言語の原文中の翻訳対象語句が含まれている自然文を検索する検索手段と、

前記原文中の各単語に対し前記原文中の翻訳対象語句との前記原文上での距離の小ささに応じた第1評価値を各々付与し、前記検索手段による検索によって抽出された自然文について、当該自然文の各単語のうち前記原文との一致単語を認識し、認識した一致単語に付与した前記第1評価値の合計値を自然文毎に演算し、演算した前記第1評価値の合計値に基づいて選択した自然文の対訳文における少なくとも前記翻訳対象語句の対訳を、前記原文中の少なくとも前記翻訳対象語句の対訳と判断する対訳判断手段と、

を含む対訳判断装置。

[6] (補正後) 前記対訳判断手段は、前記第1評価値の演算結果が同一の自然文が複数存在していた場合に、当該複数の自然文の各々について、前記自然文中の一致単語に対し前記自然文中の翻訳対象語句との前記自然文上での距離の小ささに応じた第2評価値を各々付与し、前記一致単語に付与した前記第2評価値の合計値を自然文毎に演算し、演算した前記第2評価値の合計値に基づいて選択した自然文の対訳文における少なくとも前記翻訳対象語句の対訳を、前記原文中の少なくとも前記翻訳対象語句の対訳と判断することを特徴とする請求項5記載の対訳判断装置。

[7] (補正後) 複数の単語から成る原言語の自然文を目的言語の対訳文と対応付けて複数記憶する記憶手段と、

前記記憶手段に記憶されている原言語の複数の自然文の中から、原言語の原文中の翻訳対象語句が含まれている自然文を検索する検索手段と、

前記検索手段による検索によって抽出された自然文に対し、前記原文上で前記翻訳対象語句よりも前側に存在している前側単語群のうち未探索でかつ前記翻訳対象語句との距離が最小の特定前側単語が、前記自然文上で前記翻訳対象語句又は前回の探索で抽出された前側単語の前側に所定単語数以内に存在しているか否かを探索し、前記特定前側単語が発見された場合に発見された前記特定前側単語と前記翻訳対象語句又は前回の探索で抽出された前側単語との距離の小ささに応じた第3評価値を付与することを、未探索の前側単語が前記原文上に存在しなくなる迄繰り返すと共に、前記原文上で前記翻訳対象語句よりも後側に存在している後側単語群のうちのうち未探索でかつ前記翻訳対象語句との距離が最小の特定後側単語が、前記自然文上で前記翻訳対象語句又は前回の探索で抽出された後側単語の後側に所定単語数以内に存在しているか否かを探索し、前記特定後側単語が発見された場合に発見された前記特定後側単語が前記翻訳対象語句又は前回の探索で抽出された後側単語との距離の小ささに応じた第3評価値を付与することを、未探索の後側単語が前記原文上に存在しなくなる迄繰り返し、付与した第3評価値の合計値を自然文毎に演算し、演算した前記第3評価値の合計値に基づいて選択した自然文の対訳文における少なくとも前記翻訳対象語句の対訳を、前記原文中の少なくとも前記翻訳対象語句の対訳と判断する対訳判断手段と、

を含む対訳判断装置。

[8] (補正後) 複数の単語から成る原言語の自然文を目的言語の対訳文と対応付けて複数記憶する記憶手段と、

前記記憶手段に記憶されている原言語の複数の自然文の中から、原言語の原文中の翻訳対象語句が含まれている自然文を検索する検索手段と、

前記検索手段による検索によって抽出された自然文に対し、前記原文上で前記翻訳対象語句よりも前側に存在している第1の前側単語群のうち未探索でかつ前記翻訳対象語句との距離が最小の特定前側単語が、前記自然文上で前記翻訳対象語句又は前回の探索で抽出された前側単語の前側に所定単語数以内に存在しているか否か

を探索し、前記特定前側単語が発見された場合に前記翻訳対象語句又は前回の探索で抽出された前側単語との距離に対応する前側についての第1の距離別単語数をカウントアップすることを、未探索の前側単語が前記第1の前側単語群の中に存在しなくなる迄繰り返すと共に、前記検索手段による検索によって抽出された自然文に対し、前記原文上で前記翻訳対象語句よりも後側に存在している第1の後側単語群のうち未探索でかつ前記翻訳対象語句との距離が最小の特定後側単語が、前記自然文上で前記翻訳対象語句又は前回の探索で抽出された後側単語の後側に所定単語数以内に存在しているか否かを探索し、前記特定後側単語が発見された場合に前記翻訳対象語句又は前回の探索で抽出された後側単語との距離に対応する後側についての第1の距離別単語数をカウントアップすることを、未探索の後側単語が前記第1の後側単語群の中に存在しなくなる迄繰り返し、前側についての第1の距離別単語数及び後側についての第1の距離別単語数から求めた総合評価に基づいて選択した自然文の対訳文における少なくとも前記翻訳対象語句の対訳を、前記原文中の少なくとも前記翻訳対象語句の対訳と判断する対訳判断手段と、

を含む対訳判断装置。

[9] (補正後) 前記対訳判断手段は、前記検索手段による検索によって抽出された自然文に対し、前記原文上で前記翻訳対象語句よりも前側に存在しておりかつ前記翻訳対象語句と隣り合っている前側単語を除外した第2の前側単語群のうち未探索でかつ前記翻訳対象語句との距離が最小の特定前側単語が、前記自然文上で前記翻訳対象語句又は前回の探索で抽出された前側単語の前側に所定単語数以内に存在しているか否かを探索し、前記特定前側単語が発見された場合に発見された特定前側単語と前記翻訳対象語句又は前回の探索で抽出された前側単語との距離に1又は前回の探索で発見された前側単語について求めた距離を加算した距離に対応する前側についての第2の距離別単語数をカウントアップすることを、未探索の前側単語が前記原文上に存在しなくなる迄繰り返し、前記前側についての第1の距離別単語数と前記前側についての第2の距離別単語数のうち単語間距離がより短いとの結果を示している方を最終評価として選択すると共に、前記検索手段による検索によって抽出された自然文に対し、前記原文上で前記翻訳対象語句よりも後側に存在しておりか

つ前記翻訳対象語句と隣り合っている後側単語を除外した第2の後側単語群のうち未探索でかつ前記翻訳対象語句との距離が最小の特定後側単語が、前記自然文上で前記翻訳対象語句又は前回の探索で抽出された後側単語の後側に所定単語数以内に存在しているか否かを探査し、前記特定後側単語が発見された場合に発見された特定後側単語と前記翻訳対象語句又は前回の探索で抽出された後側単語との距離に1又は前回の探索で発見された後側単語について求めた距離を加算した距離に対応する後側についての第2の距離別単語数をカウントアップすることを、未探索の後側単語が前記原文上に存在しなくなる迄繰り返し、前記後側についての第1の距離別単語数と前記後側についての第2の距離別単語数のうち単語間距離がより短いとの結果を示している方を最終評価として選択し、前側及び後側についての最終評価を加算集計して総合評価を求め、求めた総合評価において単語間距離がより短いとの結果を示している自然文の対訳文における少なくとも前記翻訳対象語句の対訳を、前記原文中の少なくとも前記翻訳対象語句の対訳と判断することを特徴とする請求項8記載の対訳判断装置。

[10] (補正後) 目的言語の対訳文と対応付けて記憶手段に複数記憶されている、各々複数の単語から成る原言語の自然文の中から、原言語の原文中の翻訳対象語句が含まれている自然文を検索する第1のステップ、

前記第1のステップにおける検索によって抽出された自然文に基づいて、前記翻訳対象語句と原言語の同一の文中に出現する頻度の高い頻出語句を認識する第2のステップ、

前記第1のステップにおける検索によって抽出された自然文のうち、前記第2のステップで認識されかつ前記原文中に存在する特定の頻出語句と、前記翻訳対象語句が各々含まれている自然文の対訳文を参照することで、前記翻訳対象語句及び前記特定の頻出語句が各々含まれる自然文の対訳文における翻訳対象語句についての高頻度の対訳を認識し、認識した高頻度の対訳を、前記原文中の前記翻訳対象語句の対訳と判断する第3のステップ

を含む対訳判断方法。

[11] (補正後) 目的言語の対訳文と対応付けて記憶手段に複数記憶されている、各々複

数の単語から成る原言語の自然文の中から、原言語の原文中の翻訳対象語句が含まれている自然文を検索する第1のステップ、

前記原文中に存在しかつ前記第1のステップにおける検索によって抽出された自然文中に含まれていない着目語句と代替可能な代替語句を判断する第2のステップ

、
前記第1のステップにおける検索によって抽出された自然文のうち、前記第2のステップで判断した代替語句及び前記翻訳対象語句が各々含まれている自然文の対訳文における少なくとも前記翻訳対象語句の対訳を、前記原文中の少なくとも前記翻訳対象語句の対訳と判断する第3のステップ

を含む対訳判断方法。

[12] (補正後) 目的言語の対訳文と対応付けて記憶手段に複数記憶されている、各々複数の単語から成る原言語の自然文の中から、原言語の原文中の翻訳対象語句が含まれている自然文を検索する第1のステップ、

前記原文中の各単語に対し前記原文中の翻訳対象語句との前記原文上の距離の小ささに応じた第1評価値を各々付与し、前記第1のステップにおける検索によって抽出された自然文について、当該自然文の各単語のうち前記原文との一致単語を認識し、認識した一致単語に付与した前記第1評価値の合計値を自然文毎に演算し、演算した前記第1評価値の合計値に基づいて選択した自然文の対訳文における少なくとも前記翻訳対象語句の対訳を、前記原文中の少なくとも前記翻訳対象語句の対訳と判断する第2のステップ

を含む対訳判断方法。

[13] (補正後) 目的言語の対訳文と対応付けて記憶手段に複数記憶されている、各々複数の単語から成る原言語の自然文の中から、原言語の原文中の翻訳対象語句が含まれている自然文を検索する第1のステップ、

前記第1のステップにおける検索によって抽出された自然文に対し、前記原文上で前記翻訳対象語句よりも前側に存在している前側単語群のうち未探索でかつ前記翻訳対象語句との距離が最小の特定前側単語が、前記自然文上で前記翻訳対象語句又は前回の探索で抽出された前側単語の前側に所定単語数以内に存在しているか否

かを探索し、前記特定前側単語が発見された場合に発見された前記特定前側単語と前記翻訳対象語句又は前回の探索で抽出された前側単語との距離の小ささに応じた第3評価値を付与することを、未探索の前側単語が前記原文上に存在しなくなる迄繰り返すと共に、前記原文上で前記翻訳対象語句よりも後側に存在している後側単語群のうちのうち未探索でかつ前記翻訳対象語句との距離が最小の特定後側単語が、前記自然文上で前記翻訳対象語句又は前回の探索で抽出された後側単語の後側に所定単語数以内に存在しているか否かを探索し、前記特定後側単語が発見された場合に発見された前記特定後側単語が前記翻訳対象語句又は前回の探索で抽出された後側単語との距離の小ささに応じた第3評価値を付与することを、未探索の後側単語が前記原文上に存在しなくなる迄繰り返し、付与した第3評価値の合計値を自然文毎に演算し、演算した前記第3評価値の合計値に基づいて選択した自然文の対訳文における少なくとも前記翻訳対象語句の対訳を、前記原文中の少なくとも前記翻訳対象語句の対訳と判断する第2のステップ

を含む対訳判断方法。

[14] (補正後) 目的言語の対訳文と対応付けて記憶手段に複数記憶されている、各々複数の単語から成る原言語の自然文の中から、原言語の原文中の翻訳対象語句が含まれている自然文を検索する第1のステップ、

前記第1のステップにおける検索によって抽出された自然文に対し、前記原文上で前記翻訳対象語句よりも前側に存在している第1の前側単語群のうち未探索でかつ前記翻訳対象語句との距離が最小の特定前側単語が、前記自然文上で前記翻訳対象語句又は前回の探索で抽出された前側単語の前側に所定単語数以内に存在しているか否かを探索し、前記特定前側単語が発見された場合に前記翻訳対象語句又は前回の探索で抽出された前側単語との距離に対応する前側についての第1の距離別単語数をカウントアップすることを、未探索の前側単語が前記第1の前側単語群の中に存在しなくなる迄繰り返すと共に、前記第1のステップにおける検索によって抽出された自然文に対し、前記原文上で前記翻訳対象語句よりも後側に存在している第1の後側単語群のうち未探索でかつ前記翻訳対象語句との距離が最小の特定後側単語が、前記自然文上で前記翻訳対象語句又は前回の探索で抽出された後側単語の

後側に所定単語数以内に存在しているか否かを探索し、前記特定後側単語が発見された場合に前記翻訳対象語句又は前回の探索で抽出された後側単語との距離に対応する後側についての第1の距離別単語数をカウントアップすることを、未探索の後側単語が前記第1の後側単語群の中に存在しなくなる迄繰り返し、前側についての第1の距離別単語数及び後側についての第1の距離別単語数に基づいて選択した自然文の対訳文における少なくとも前記翻訳対象語句の対訳を、前記原文中の少なくとも前記翻訳対象語句の対訳と判断する第2のステップ

を含む対訳判断方法。

[15] (補正後) 複数の単語から成る原言語の自然文を目的言語の対訳文と対応付けて複数記憶する記憶手段と接続されたコンピュータを、

前記記憶手段に記憶されている原言語の複数の自然文の中から、原言語の原文中の翻訳対象語句が含まれている自然文を検索する検索手段、

前記検索手段による検索によって抽出された自然文に基づいて、前記翻訳対象語句と原言語の同一の文中に出現する頻度の高い頻出語句を認識する認識手段、

及び、前記検索手段による検索によって抽出された自然文のうち、前記認識手段によって認識されかつ前記原文中に存在する特定の頻出語句と、前記翻訳対象語句が各々含まれている自然文の対訳文を参照することで、前記翻訳対象語句及び前記特定の頻出語句が各々含まれる自然文の対訳文における翻訳対象語句についての高頻度の対訳を認識し、認識した高頻度の対訳を、前記原文中の前記翻訳対象語句の対訳と判断する対訳判断手段

として機能させるプログラム。

[16] (補正後) 複数の単語から成る原言語の自然文を目的言語の対訳文と対応付けて複数記憶する記憶手段と接続されたコンピュータを、

前記記憶手段に記憶されている原言語の複数の自然文の中から、原言語の原文中の翻訳対象語句が含まれている自然文を検索する検索手段、

前記原文中に存在しかつ前記検索手段による検索によって抽出された自然文中に含まれていない着目語句と代替可能な代替語句を判断する判断手段、

及び、前記検索手段による検索によって抽出された自然文のうち、前記判断手段

によって判断された代替語句及び前記翻訳対象語句が各々含まれている自然文の対訳文における少なくとも前記翻訳対象語句の対訳を、前記原文中の少なくとも前記翻訳対象語句の対訳と判断する対訳判断手段

として機能させるプログラム。

[17] (補正後) 複数の単語から成る原言語の自然文を目的言語の対訳文と対応付けて複数記憶する記憶手段と接続されたコンピュータを、

前記記憶手段に記憶されている原言語の複数の自然文の中から、原言語の原文中の翻訳対象語句が含まれている自然文を検索する検索手段、

及び、前記原文中の各単語に対し前記原文中の翻訳対象語句との前記原文上での距離の小ささに応じた第1評価値を各々付与し、前記検索手段による検索によって抽出された自然文について、当該自然文の各単語のうち前記原文との一致単語を認識し、認識した一致単語に付与した前記第1評価値の合計値を自然文毎に演算し、演算した前記第1評価値の合計値に基づいて選択した自然文の対訳文における少なくとも前記翻訳対象語句の対訳を、前記原文中の少なくとも前記翻訳対象語句の対訳と判断する対訳判断手段

として機能させるプログラム。

[18] (補正後) 複数の単語から成る原言語の自然文を目的言語の対訳文と対応付けて複数記憶する記憶手段と接続されたコンピュータを、

前記記憶手段に記憶されている原言語の複数の自然文の中から、原言語の原文中の翻訳対象語句が含まれている自然文を検索する検索手段、

及び、前記検索手段による検索によって抽出された自然文に対し、前記原文上で前記翻訳対象語句よりも前側に存在している前側単語群のうち未探索でかつ前記翻訳対象語句との距離が最小の特定前側単語が、前記自然文上で前記翻訳対象語句又は前回の探索で抽出された前側単語の前側に所定単語数以内に存在しているか否かを探索し、前記特定前側単語が発見された場合に発見された前記特定前側単語と前記翻訳対象語句又は前回の探索で抽出された前側単語との距離の小ささに応じた第3評価値を付与することを、未探索の前側単語が前記原文上に存在しなくなる迄繰り返すと共に、前記原文上で前記翻訳対象語句よりも後側に存在している後側単語

群のうちのうち未探索でかつ前記翻訳対象語句との距離が最小の特定後側単語が、前記自然文上で前記翻訳対象語句又は前回の探索で抽出された後側単語の後側に所定単語数以内に存在しているか否かを探索し、前記特定後側単語が発見された場合に発見された前記特定後側単語が前記翻訳対象語句又は前回の探索で抽出された後側単語との距離の小ささに応じた第3評価値を付与することを、未探索の後側単語が前記原文上に存在しなくなる迄繰り返し、付与した第3評価値の合計値を自然文毎に演算し、演算した前記第3評価値の合計値に基づいて選択した自然文の対訳文における少なくとも前記翻訳対象語句の対訳を、前記原文中の少なくとも前記翻訳対象語句の対訳と判断する対訳判断手段

として機能させるプログラム。

[19] (補正後) 複数の単語から成る原言語の自然文を目的言語の対訳文と対応付けて複数記憶する記憶手段と接続されたコンピュータを、

前記記憶手段に記憶されている原言語の複数の自然文の中から、原言語の原文中の翻訳対象語句が含まれている自然文を検索する検索手段、

及び、前記検索手段による検索によって抽出された自然文に対し、前記原文上で前記翻訳対象語句よりも前側に存在している第1の前側単語群のうち未探索でかつ前記翻訳対象語句との距離が最小の特定前側単語が、前記自然文上で前記翻訳対象語句又は前回の探索で抽出された前側単語の前側に所定単語数以内に存在しているか否かを探索し、前記特定前側単語が発見された場合に前記翻訳対象語句又は前回の探索で抽出された前側単語との距離に対応する前側についての第1の距離別単語数をカウントアップすることを、未探索の前側単語が前記第1の前側単語群の中には存在しなくなる迄繰り返すと共に、前記検索手段による検索によって抽出された自然文に対し、前記原文上で前記翻訳対象語句よりも後側に存在している第1の後側単語群のうち未探索でかつ前記翻訳対象語句との距離が最小の特定後側単語が、前記自然文上で前記翻訳対象語句又は前回の探索で抽出された後側単語の後側に所定単語数以内に存在しているか否かを探索し、前記特定後側単語が発見された場合に前記翻訳対象語句又は前回の探索で抽出された後側単語との距離に対応する後側についての第1の距離別単語数をカウントアップすることを、未探索の後側単語が前

記第1の後側単語群の中に存在しなくなる迄繰り返し、前側についての第1の距離別単語数及び後側についての第1の距離別単語数に基づいて選択した自然文の対訳文における少なくとも前記翻訳対象語句の対訳を、前記原文中の少なくとも前記翻訳対象語句の対訳と判断する対訳判断手段

として機能させるプログラム。